Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG

aMP: Deep learning framework to characterize meiosis progression timeline in tetraploid *Arabidopsis thaliana*

PhD Thesis

by

Ankit Ghosh

Start date:	January 1, 2019
End date:	February 22, 2024
Advisor and 1st Examiner:	Prof. Dr. Arp Schnittger
2nd Examiner:	Prof. Dr. Sören Laue
Defense date:	October 1, 2024

Abstract

In plant science, the detailed examination of meiosis progression is hampered by the labour-intensive process of manual image analysis. Our research presents a novel, deep learning-based automation pipeline that significantly streamlines the quantification of meiotic timelines through the analysis of live-imaging videos. This innovative approach employs segmentation models to stabilize video frames, refines localization techniques to precisely identify individual meiocytes, and adopts a restricted space motion-inspired tracking methodology that effectively reduces computation time and improves tracking accuracy. Crucially, our framework distinguishes itself by generating Z-normalized staging pathways, enabling the construction of a piece-wise timeline of meiotic progression. This is achieved through a meticulously curated landmarking scheme, which our results confirm aligns with established meiosis timelines in both wild-type and heat-shocked *Arabidopsis thaliana*.

Our study ventures beyond the diploid paradigm, extending the application of our high-throughput pipeline to tetraploid variants. The analyses disclose that while tetraploids exhibit comparable meiosis-I timelines to their diploid counterparts, a pronounced prolongation characterizes their meiosis-II stages. Furthermore, the systematic examination of *tcx5;6* mutants and ATM gene insertions in tetraploids provides a quantitative view of the temporal dynamics in meiotic progression, highlighting the potential for chromosomal behaviour and genetic regulation to modulate meiotic efficiency.

By integrating a convolutional neural network (CNN) based methodology with our modular pipeline, we deliver a transformative tool for meiosis analysis. Our work is not only restricted to timeline analysis, but the modular approach shows ability in different segmentation tasks from basics like pollen counting to more structured like DNA double-strand break and BiFC quantization.

Statement

Hereby, I state that this work has been prepared by myself and with the help referred to within this thesis. I would like to additionally state that I have utilized large language models (Llama65b from META on a personal computer and chatGPT3.5 from OpenAI through their API service) for grammatical correction and language coherence. Additionally, I have used Grammarly for sentence fragmentation, rephrasing, and word choice.

Hamburg, 22nd February 2024

Eidesstattliche Versicherung
Declaration on oath

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

I hereby declare, on oath, that I have written the present dissertation on my own and have not used other than the acknowledged resources and aids.

Ort, Datum Unterschrift

Acknowledgement

I express my deepest gratitude to my advisor, Prof. Dr. Arp Schnittger, for the unwavering support and guidance throughout this research. Your expertise and insights have been invaluable to my work. Our discussions have profoundly shaped my understanding and approach to this field. I am and will remain deeply indebted to you for your patient mentorship and for believing in my potential, even in the face of our differences in opinion. I feel strongly that your faith in my work aimed to achieve a publication of the highest quality, which I consider imminent.

My sincere thanks also go to my thesis committee members, Prof. Dr. Sören Laue, Prof. Philip Porada, and Magdalena Weingartner, for their readiness to evaluate my work and contribute to the seamless completion of my project. Their dedication to thoroughly assess my academic work is highly appreciated.

I am indebted to the Department of Developmental Biology at the Universität Hamburg for providing the essential facilities and resources for my research. I appreciate my fellow lab members and colleagues, especially Maria, Franzi, Óscar, Lucas, Yuki, Hasibe, Joke, and Max, for their collaboration, encouragement, and the intellectually stimulating environment they have fostered. I am particularly grateful to Zhijian for his selfless assistance in benchmarking my work. My biological journey owes much to you all, and I cannot overstate your role in keeping me committed to my research and presenting me with continual challenges.

Special thanks are due to Maren Röper and Gaetan Pochon, whose support was crucial during my research and initial days in the lab. Their patience and expertise in managing day-to-day lab activities have been indispensable.

I want to acknowledge the Human Frontier Science Program (HFSP) for providing the financial support that made my research possible. I am also thankful to the administrative and staff members of Universität Hamburg for their assistance with the logistical aspects of my thesis work.

On a personal note, I am indebted to my mother for her for their love, moral support, and patience. I must mention Mohammad Salman, who encouraged my application to the position, and my father, who insightfully advised that embracing this opportunity was the way forward. My mother, brother, and uncle have been constant pillars of mental support, helping me set realistic expectations for a PhD journey. To Megha, who stood by me during my most challenging times, your unwavering support has been my anchor, and I am hopeful for brighter days ahead.

Lastly, I extend my appreciation to everyone who, directly or indirectly, contributed to the

completion of this thesis. Your collective wisdom and encouragement have been the compass that guided me through this academic voyage, preserving my sanity. I cherish all the dinners, get-togethers, football matches, and the diverse cuisines we shared - the taste of Seiji's Okonomiyaki lingers memorably.

Contents

List of Figures

List of Tables

List of Symbols

c	Cell shape state.
n	Nucleus position state.
o	Nucleolus position state.
r	Chromatin state based on REC8 expression.
m	Microtubular array state.
siv	State indicative vector.
k	Meiocyte at a timepoint.
$\mathbb{V}$	The original video sequence with frames $\mathbb{V}_t$ for $t = 0, \ldots, T$
$\mathbb{M}$	Sequence of segmented masks corresponding to anther regions, with masks $\mathbb{M}_t$
$C(\mathbb{M}_t)$	Centroid of the anther region in the mask $\mathbb{M}_t$ for each frame t
ΔC_t	Displacement vector of anther's centroid from first frame to current frame t
τ	Transformation function combining translational and rotational adjustments
$\vec{T}_t$	Translation matrix for frame t
R_t	Rotation matrix for frame t
$\mathbb{V}'_t$	Transformed video frame after applying inverse transformation τ_t^{-1}
$\mathbb{M}'_t$	Transformed mask after applying inverse transformation τ_t^{-1}
$P_{i,j}$	Position of contour centroid j in frame i
$\vec{T}_{i,j}$	Translation vector for contour j from frame i to the first frame
$R_{i,j}$	Rotation matrix for contour j in frame i
$D_{i,j}$	Displacement of contour centroid j in frame i after stabilization
$\theta_{i,j}$	Angle of rotation for alignment of contour j in frame i
$\mathbb{V}'$	Stabilized video frames.
t	Time point index in a video sequence, ranging from 0 to T.
T	Final time point in a video sequence.
$\mathbb{V}'_t$	Individual stabilized video frame at time point t.
$C_{V_{t=0,\ldots,T}}$	Set of localized meiocytes across the video frames.
$\mathbb{C}_t$	Cropped image segments corresponding to localized meiocytes in frame $\mathbb{V}'_t$.
B	Set of bounding boxes.
P	Associated confidence scores for bounding boxes.
CT	Predefined confidence threshold.
B''	Filtered set of bounding boxes after object filtering.
P''	Corresponding confidence scores for filtered bounding boxes.
b_i	Bounding box with the highest confidence score.
N_t	IoU threshold for non-maximum suppression.
B'	Final, non-redundant set of bounding boxes after non-maximum suppression.
$w(C)$	Width of the bounding box for contour C.

$h(C)$	Height of the bounding box for contour C.
$\theta(C)$	Orientation angle of the bounding box for contour C.
ϵ	Distance threshold used in the RSS.
$d_{t,t-1}$	Euclidean distance between meiocyte detections in consecutive frames.
$\mathbf{pO}(\cdot)$	Function calculating the centroid of a meiocyte's bounding box.
$\text{NCC}_{i,j}$	NCC score for two datapoints
$\text{IoU}_{i,j}$	IoU score for two datapoints
α, β	Weighting factors for combining NCC and IoU results.
$T^i_{\mathbb{V}_t}$	Tracking update rule for meiocyte i in frame t.
$d_{t,i}$	Distance between predicted and ground truth position of an object.
c_t	Number of correctly matched pairs at time t.
$\mathbb{S}'_t$	Initial state vector predicted for meiocytes at time t
$\mathbb{S}''_t$	Final state vector after knowledge graph validation for meiocytes at time t
$\mathbb{C}_t$	Cropped image segments of meiocytes at time t
A	Adjacency matrix representing the knowledge graph
$\text{Valid}(\cdot, \cdot)$	Function to validate transitions against the knowledge graph
$\mathbb{S}'$	State vector output by the classifier
$\mathbb{S}''$	Adjusted state vector after knowledge graph post-processing
ϵ	Threshold used in knowledge graph validation
G	The knowledge graph hline
siv	State Indicative Vector
$L(siv_k, siv_j)$	Distance between state-indicative vectors
ns_k	Neighboring state of siv_k
μ_{siv_k}	Mean occurrence of all neighboring states of siv_k
σ_{siv_k}	Standard deviation of the occurrence of all neighboring states of siv_k
Z_{siv_k}	Z-score for each state-indicative vector siv_k
A_l	Landmark states
T_l	Transition time between landmarks
c_{siv}	Count of occurrences of state siv in the dataset
$c_{ns_{kj}}$	Count of occurrences of neighboring state ns_{kj}
T_s	Sampling time of the video $\mathbb{V}$.
a_l	Meiosis transition stages.
t_{A_l}	Time spent in transition stages.

Chapter 1

Introduction

Meiosis, an intricate cell division process, involves two sequential chromosomal segregation, namely meiosis I and meiosis II, following a single round of DNA replication, resulting in a halving of the DNA content. This reduction is crucial for sexually reproducing organisms as it maintains the genome size from generation to generation, ultimately facilitating diploid restoration post-gamete fusion. Additionally, meiosis plays a vital role in generating genetic diversity. During early prophase I, homologous chromosomes undergo reciprocal DNA exchanges via crossing-over, forming novel genetic alleles. Furthermore, the random separation of homologous chromosome pairs at the end of meiosis-I generates diverse yet complete chromosome sets in daughter cells.

Prusicki et al. have previously pioneered a live cell imaging system using a confocal microscope to observe the dynamic characteristics of meiosis [3]. This robust approach facilitated the in vivo examination of floral meiosis over its entire course. It provided images from deep tissue layers to describe the progression of meiosis using a novel landmarking system. The landmark system relies on determining five cellular features (parameters): cell shape (CS), intracellular nucleus position (NP), nucleolus position (NoP) within the nucleus, chromatin (RC) condensation level, and the pattern of the microtubule (MT) cytoskeleton; see Fig. 2.4. By applying this approach to the model plant wild-type A.thaliana, our group was able to identify clearly defined states of meiosis characterized by specific parameter configurations, enabling both qualitative and quantitative analysis of the meiotic process.

1.1 Task description

Identification of the landmarks required manual tracking of meiocytes within microscopy videos and staging of the meiocytes, i.e., manual annotation of video frames. Such manual analysis is highly time-consuming and error-prone due to variability between and within observers. Motivated by the author's work and the promising capabilities of current deep learning systems in computer vision, the present work aimed to automate the analysis, significantly reducing the related effort and accelerating the research process.

The automation confronts several challenges due to the nature of the meiosis-related data, especially the movement of anthers containing the meiocytes, variation in contrast of videos and

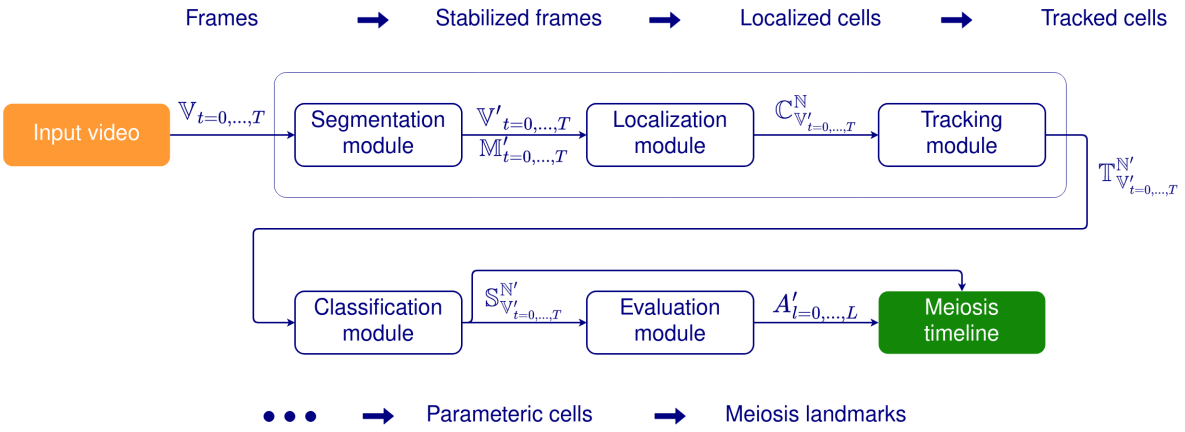

Figure 1.1: ***Overview of the modular aMP framework***
The input is the confocal live imaging video $\mathbb{V}$ to be analyzed. Within the video, anther regions containing meiocytes are segmented and anther motion is compensated, resulting in motion-compensated video $\mathbb{V}'$) and corresponding anther masks $\mathbb{M}'$. Within the individual frames $t = 0, \dots, T$, the meiocytes are localized ($\mathbb{C}_{\mathbb{V}'_t}$: set of meiocytes in frame $\mathbb{V}'_t$) and tracked over time ($\mathbb{T}_{\mathbb{V}'_t}$). The meiocytes are finally classified according to their meiosis state. The resulting state information $\mathbb{S}_{\mathbb{V}'_t}$ of the video (and similar videos of the underlying imaging experiment) is used to derive frequently occurring states, meiosis landmarks $A_{l=0,\dots,L}$, that characterize the meiosis landscape.

video frames, occlusion effects, and classification ambiguity. Additionally, meiocyte morphological changes can occur gradually or abruptly, depending upon the frequency of live imaging. These challenges require the development of sophisticated algorithms and models for the individual sub-tasks. Here, we propose a modular pipeline for the automated analysis of meiotic progression (aMP-kit), using convolutional neural networks (CNNs) as the primary computational tool.

The pipeline and its blocks are sketched in Fig. 1.1 and detailed in chapter 3 & 4. The main modules are:

1. a **segmentation block** that segments the anther regions relevant for subsequent meiocyte localization and compensates for anther movements between frames,
2. a **localization block** that identifies the meiocytes within the anthers,
3. a **tracking block** that tracks the meiocytes over time,
4. a **classification block** that classifies the state of the meiocytes according to the five cellular features identified by Prusicki et al. and the corresponding feature expressions and
5. an **evaluation block** that identifies meiosis landmark states based on the classification states for the data collective considered.

1.2 Outline of the project thesis

The thesis has been divided into three major parts - introduction to the A. thaliana as a model plant, meiotic cell division and the data characteristics; tracking of meiocytes through image

stabilization and object localization, staging of meiocytes and subsequently; and finally, the evaluation processes required to study different cellular dynamics.

Chapter 2 will discuss about the confocal live-imaging meiosis dataset. We will discuss the biological prior knowledge derived from the dataset and how they are incorporated into pipeline automation. Furthermore, we will discuss the preprocessing and augmentation techniques used to enhance the dataset's quality and diversity. We will also describe the additional datasets used to check the automated pipeline's sanity.

Chapter 3 will focus on the methods and the algorithm developed for each of the blocks of the modular pipeline. We will introduce focal stabilization, meiocyte tracking through localization, meiocyte staging, and timeline generation. We will discuss the concept of convolutional neural networks, feature extraction and block-specific post-extraction processes. We will also describe the principles applied to evaluate each block. We will also enlist the involved (hyper)parameters and other important implementation details.

Chapter 4 will guide you through the network-optimized results and corresponding evaluation. We will discuss the relevance of the biological outcomes obtained from the automated single-cell analysis of meiosis progression. We will discuss how the insights from the study contribute to a better understanding of meiotic processes in A. thaliana and potentially other related species. Additionally, we will explore the broader applications of the analytical pipeline beyond meiosis research. This will include discussing other areas of biological research where the automated single-cell analysis approach can be adapted and its potential to extend to various use cases. The chapter will highlight the significance of the findings and their implications for advancing biological studies.

Chapter 5 delves into the versatile applications of our framework beyond its current scope, highlighting its adaptability to various segmentation tasks within analytical plant science. This section underscores the significance of automating critical processes such as foci quantification, biomolecular fluorescence complementation (BiFC) analysis, and pollen counting. These examples serve to illustrate the promising avenues for future research, focusing on the development of an inclusive analytical framework. This envisioned framework aims to integrate Large Language Models and Generative AI within the realm of computer vision, thereby expanding the frontiers of plant science research through enhanced computational techniques

Chapter 2

Live cell imaging dataset of *Arabidopsis thaliana*

This chapter explores the foundational elements necessary to analyse meiotic progression in Arabidopsis thaliana. Central to this pursuit is the comprehensive understanding and preparation of the datasets utilized in our study. These datasets, derived from meticulous live cell imaging, serve as the bedrock upon which our computational models will be built and tested.

Data quality and preparation are as critical as the analytical methods in biological research, particularly in studying complex processes like meiosis. Thus, we begin by detailing the acquiring and preparing the Arabidopsis thaliana datasets. This includes thoroughly discussing the preprocessing steps to ensure that the data is accurate and representative of the biological phenomena under study and formatted appropriately for subsequent analysis using state-of-the-art AI techniques.

We then delve into the strategies employed for data augmentation, a crucial step in enhancing the robustness and efficacy of the machine learning models we intend to develop. Data augmentation simulates a broader range of conditions and variations the models might encounter, enriching the training process and enhancing model generalizability.

Furthermore, we address the strategies for splitting the data into training, validation, and testing sets. This segmentation is vital for unbiased model training and rigorous performance evaluation. It ensures that the models are tested against data not encountered during training, providing a realistic assessment of their predictive capabilities.

Finally, this chapter sets the stage for applying AI and machine learning methodologies. The transition from raw, preprocessed data to the utilization of sophisticated computational models marks a significant phase in our research. Here, we begin to transform the rich datasets into meaningful scientific insights, leveraging the power of AI to unravel the complexities of meiotic progression in various strains of Arabidopsis thaliana, laying the foundation for the AI-driven exploration central to achieving our research objectives.

2.1 Arabidopsis thaliana: an overview

Arabidopsis thaliana, commonly called thale cress, is a small flowering plant that has become a model organism in plant genetics and molecular biology [4]. Originating from Eurasia, this plant's significance in research parallels that of the fruit fly and other model organisms in zoological studies. Its small genome was fully sequenced in 2000, marking a significant milestone in plant biology and positioning it as a prime candidate for genetic studies [5]. The short 6-8 weeks life cycle allows for rapid generational studies, an advantageous trait for genetic and developmental biology research [6]. Arabidopsis's amenability to genetic manipulation further enhances its value in the research community, allowing for extensive studies into gene function and regulation [7]. Its status as a model organism is bolstered by the wide range of genetic and molecular tools available, including extensive mutant collections and genomic databases, making it an invaluable resource for understanding broader biological processes in plants [8]. Arabidopsis thaliana has been prominently featured in seminal scientific literature, playing a central role in the progression of our comprehension of plant genetics and molecular biology. Meyerowitz et al. (1989) laid the groundwork for establishing Arabidopsis as a model organism by providing detailed insights into its genetic and developmental attributes. Following the complete sequencing of the Arabidopsis genome, this genomic blueprint has served as a pivotal resource, extensively employed by researchers to decipher intricate gene networks and regulatory pathways. The study conducted by Alonso et al. (2003) serves as an exemplary demonstration of Arabidopsis' utility in functional genomics, highlighting the potency of genetic modification for the precise dissection of gene function and interactions. Moreover, Sessions et al. (2002) demonstrated the significance of extensive mutant collections, enabling researchers to scrutinize specific gene functions with meticulous precision, thereby making substantial contributions to the broader understanding of plant biology. These landmark contributions affirm Arabidopsis thaliana's standing as a preeminent model organism, validating its indispensable role in advancing biological research.

Specific to the context of this thesis, Arabidopsis thaliana's suitability for live cell imaging, particularly in studying meiosis, is noteworthy. The plant's clear cellular structures and relatively simple developmental processes provide an excellent framework for observing meiotic events in real-time. This is particularly relevant in exploring the genetic and phenotypic variations in tetraploid variants and ATM gene-inserted tetraploid variants of Arabidopsis thaliana, which are central to this study.

2.2 Significance of Arabidopsis in studying meiosis progression timeline

Building upon the foundational role of *Arabidopsis thaliana* in plant genetics, its application extends to studying the temporal course of meiotic progression, particularly in various mutants. Meiotic progression analysis involves studying the stages of meiosis, a cell division mechanism essential for sexual reproduction. Understanding the mechanisms governing genetic diversity generation and perpetuation is crucial, and Arabidopsis thaliana's genetic adaptability and extensive research resources make it an ideal candidate for such analyses. The prototypical Arabidopsis serves as a reference model for standard meiotic processes. At the same time, tetraploid Arabidopsis, with its duplicated chromosomal set, offers insights into meiotic mechanisms under polyploid conditions, vital for understanding species evolution and plant breed-

ing methodologies. Mutant strains like *atm*(ATAXIA TELANGIECTASIA MUTATED) and *tam* (TARDY ASYNCHRONOUS MEIOSIS) show alteration for the meiotic progression time course, shedding light on the genetic controls of meiosis. The *tcx* mutant (Tesmin/TSO1-like CXC) provides a unique perspective on chromosomal dynamics during meiosis, particularly in synapsis and recombination events. These variants enrich our understanding of meiotic variability and genetic regulation in *Arabidopsis thaliana*, demonstrating its versatility as a model organism.

Significant advances have been made in understanding the meiotic progression timeline in wild-type *Arabidopsis thaliana*. Armstrong (2013) provided a comprehensive analysis of meiotic progression, offering critical insights into various stages of meiosis [9]. Yang et al. (2013) pioneered techniques for visualising meiotic proteins, which are vital for studying meiotic processes at the molecular level [10]. Chen and Retzel (2013) contributed by analyzing the meiotic transcriptome, shedding light on genetic regulation during meiosis [11]. Armstrong (2013) also employed cytogenetic analysis to understand chromosome behaviour during meiosis further [12]. Investigation by Liu and Qu (2008) and Mercier and Grelon (2008) explored the role of cell cycle mutants and gene discovery in meiosis over the past decade, respectively, enhancing our understanding of the genetic control and evolution of meiosis in *Arabidopsis* [1, 13].

Studies on meiotic progression in *Arabidopsis thaliana* have provided extensive insights into this fundamental biological process. Sanchez-Moran et al. (2008) and Jones et al. (2003) explored chromosome synapsis and recombination, integrating cytological and molecular approaches [14, 15]. De Muyt et al. (2009) identified new early meiotic recombination functions through a high-throughput genetic screen [16]. Naranjo and Corredor (2008) delved into nuclear architecture and chromosome dynamics in meiosis [17], while Osman et al. (2011) discussed the pathways to meiotic recombination [18]. Complementing these findings, live cell imaging research by Armstrong revealed distinct cellular states or landmarks in meiosis, offering a dynamic view of the process [9]. This dynamic view of the meiotic process complements previous research findings and adds a temporal dimension to our understanding. Combining traditional cytological and molecular approaches with cutting-edge live cell imaging techniques forms a holistic approach, allowing for a nuanced exploration of the factors influencing meiotic progression in Arabidopsis thaliana.

Prusicki et al. [3] focus on establishing a live cell imaging setup to observe male meiocytes in Arabidopsis. This method allows for the visualization of microtubules and a meiotic cohesin subunit, enabling the study of five cellular parameters: cell shape, microtubule array, nucleus position, nucleolus position, and chromatin condensation. The study identified 11 distinct cellular states or landmarks that are not randomly associated but represent convergence points during meiotic progression. This approach was used to analyze a mutant in the meiotic A-type cyclin TARDY ASYNCHRONOUS MEIOSIS (TAM), revealing both qualitative and quantitative changes in these meiotic landmarks, including the formation of ectopic spindle- or phragmoplast-like structures not attached to chromosomes.

Recent studies on tetraploids in *Arabidopsis thaliana* have significantly contributed to our comprehension of plant polyploidy. Yu et al. (2009) demonstrated the high stability of neotetraploid lines in *Arabidopsis* over consecutive generations, emphasizing the robustness of tetraploid genomes [19]. Pecinka et al. (2011) found that polyploidization in *Arabidopsis* increases meiotic recombination frequency, which is a crucial factor for genetic diversity [20]. Liu et al. (2017) discussed the rapid genomic changes associated with autopolyploidy in *Arabidopsis*, highlighting the dynamic nature of polyploid genomes [21]. Jørgensen et al. (2011) explored

gene flow in Central European Arabidopsis, including interploidal gene flow, which is vital for understanding the evolutionary dynamics of polyploids [22]. del Pozo and Ramirez-Parra (2015) provided an overview of whole genome duplications in plants, with a focus on *Arabidopsis*, elucidating the mechanisms and consequences of polyploidy [23]. Finally, Novikova et al. (2018) correlated the origin of polyploid *Arabidopsis* species to recent glaciation periods, offering insights into the evolutionary history of these plants [24].

In *Arabidopsis thaliana*, The ATM gene has been a focus of various studies, revealing its crucial roles in response to environmental stresses and in maintaining genomic stability. Culligan et al. (2006) showed that ATM and ATR play distinct and additive roles in the plant's response to ionizing radiation. Their research highlighted that ATM is essential for upregulating many genes in response to gamma-irradiation [25]. Sakamoto et al. (2009) further elucidated the role of ATM in the DNA damage response, specifically in a UVB-hypersensitive mutant of Arabidopsis, thereby underlining the importance of ATM in environmental stress responses [26]. Vespa et al. (2005) discussed the distinct contributions of ATM and ATR in chromosome end protection and telomeric DNA maintenance, indicating the multifaceted role of ATM in genomic integrity [27]. More recently, Zhao et al. (2023) demonstrated that the ATM-mediated double-strand break repair is vital for maintaining meiotic genome stability, especially under high-temperature stress, thus shedding light on the role of ATM in meiotic recombination and genome stability under stressful conditions [28]. This comprehensive exploration of ATM's functions enriches the contextual framework for understanding its significance in Arabidopsis thaliana's intricate regulatory networks.

Additionally, the molecular insights into meiotic recombination mechanisms in Arabidopsis thaliana are further enriched by including MLH3 in the investigation. MLH3 is part of the DNA mismatch repair system and is crucial in meiotic recombination. In Arabidopsis and other species, MLH3 promotes crossover events during meiosis, essential for exchanging genetic material between homologous chromosomes and generating genetic diversity. In *mlh3* mutants, the process of crossover formation during meiosis is typically impaired. This can lead to reduced fertility or sterility due to the production of gametes with aberrant chromosomal numbers or structures. Utilizing *mlh3* mutants in genetic studies provides a valuable avenue for unravelling the mechanisms governing crossover formation and delving into the genetic regulation of meiotic recombination, contributing to a deeper understanding of plant reproductive processes.

Lastly, integrating the analysis of the meiotic progression of *tcx5;6* mutants adds another layer of complexity to understanding meiotic regulation in Arabidopsis thaliana. Tesmin/TSO1-like CXC domain-containing protein 5 (TCX5) and TCX6 is part of the plant DREAM complex, a well-conserved transcriptional complex among eukaryotes that coordinates the expression of cell cycle genes. tcx5;6 double mutants show defects in pairing, synapsis and crossover formation, ultimately resulting in unbalanced chromosome segregation in *Arabidopsis* male meiocytes (unpublished data, Hasibe Tuncay). Analysis of meiotic progression in *tcx5;6* mutants will provide a better understanding of the dynamics leading to these meiotic defects.

The examination of various Arabidopsis mutants, including tetraploid and ATM gene-inserted variants, aligns closely with the primary aim of this thesis: to employ advanced AI techniques and live cell imaging for a detailed analysis of meiosis. Each aspect of Arabidopsis research discussed here provides foundational knowledge that underpins our methodological approach and analysis. This comprehensive exploration of Arabidopsis in meiotic studies enriches our understanding of plant genetics and sets the stage for our subsequent investigations. By applying cutting-edge live-imaging techniques, we seek to further elucidate the nuances of meiotic

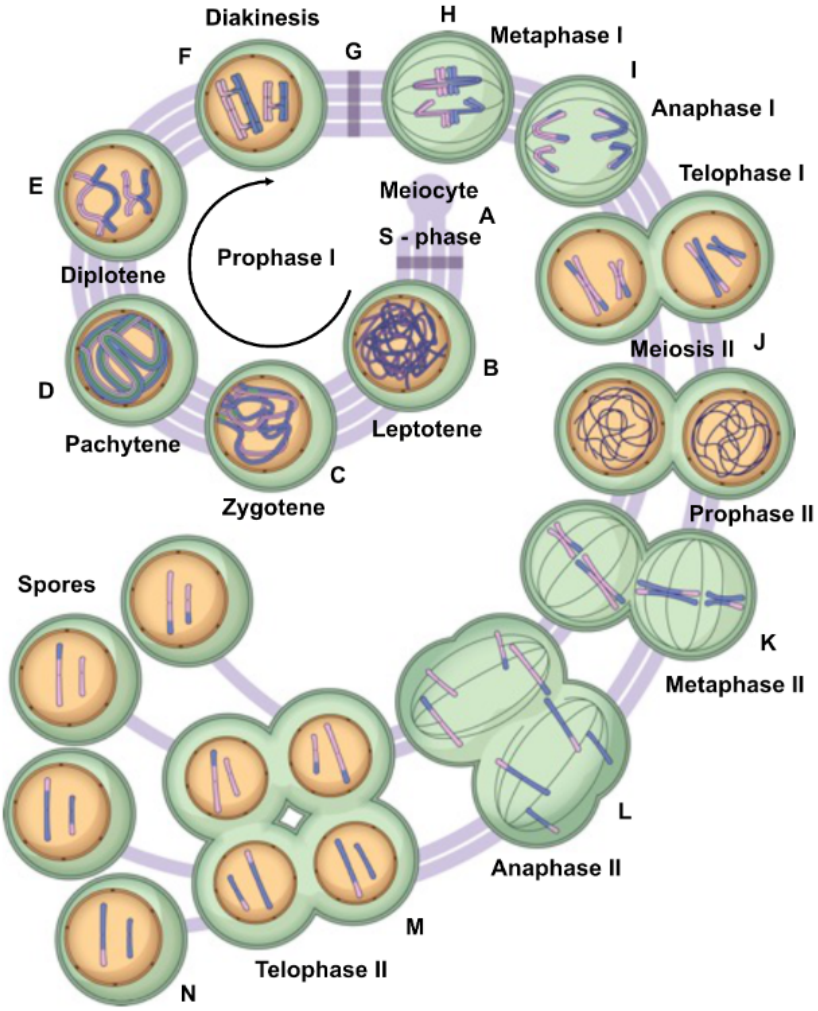

Figure 2.1: Stages of Meiosis in Plant Cells [1]
This figure depicts the sequential stages of meiosis, starting from the early prophase I through to the formation of spores. Each cell illustrates the chromosomal configurations and nuclear changes characteristic of the respective phase: A) Leptotene, B) Zygotene, C) Pachytene, D) Diplotene, E) Diakinesis, F) Metaphase I, G) Anaphase I, H) Telophase I, I) Prophase II, J) Metaphase II, K) Anaphase II, L) Telophase II, and finally M) the resulting spores. The diagram serves as a visual guide to the complex process of meiotic cell division in plants.

progression in Arabidopsis, particularly in its tetraploid variants and ATM gene-inserted lines. These insights will be pivotal in enhancing our application of AI tools for analyzing these complex biological processes. Reflecting on the imaging techniques utilized in these studies, we note the progression from traditional cytological methods to more advanced live cell imaging. This evolution mirrors our methodological journey in this thesis, where we harness state-of-the-art imaging analysis to the dynamic process of meiosis in Arabidopsis thaliana.

2.3 Meiosis in male meiocytes

Building on the established role of Arabidopsis thaliana as a key player in genetic research, we now focus on the intricate process of meiosis in this model organism. Understanding the

detailed stages of meiosis in Arabidopsis is crucial for our study, as it provides a foundational basis for our advanced imaging and AI-based analyses. As we delve into each phase of meiosis in Arabidopsis, it's important to consider how this detailed biological understanding informs and enhances our AI-driven approach. The precise identification of meiotic stages is pivotal in interpreting the live cell imaging data, especially when analyzing the more complex genetic structures found in tetraploid and ATM gene-inserted variants.

Meiosis is a fundamental biological process for producing haploid gametes from diploid cells. This process is critical for sexual reproduction in eukaryotes and contributes to genetic diversity [29]. In *Arabidopsis thaliana*, a model organism for plant genetics, meiosis has been extensively studied due to the simplicity and accessibility of the genetic tools available [30].

Male meiosis in *Arabidopsis* has been a particular research focus. Recent studies have developed methods for isolating male meiocytes, enabling detailed molecular and cytogenetic analyses of meiotic chromosomes [29]. Techniques such as spreading and fluorescence in situ hybridization have been employed to study both male and female meiocyte chromosomes in *Arabidopsis*, providing valuable insights into the cytogenetical aspects of meiosis [30]. Rapid methods for visualizing male meiotic chromosomes have also been developed, facilitating quicker and more efficient studies in this area [31].

In Arabidopsis, meiosis initiates within the floral meristem's reproductive cells, specifically within the anthers for male meiocytes and the ovules for female meiocytes. The process begins with the replication of chromosomes during the interphase stage, setting the stage for the subsequent division phases.

- **Prophase I:** Prophase I is the most complex and longest phase of meiosis, encompassing several sub-stages where crucial events of chromosome pairing, recombination, and synapses occur. In Arabidopsis thaliana, as in other organisms, Prophase I is subdivided into the following stages:

    1. **Leptotene**: This initial stage is marked by the beginning of chromosomal condensation. Chromosomes start to become visible under the light microscope as slender filaments. The process of homologous chromosome searching commences, setting the stage for synapsis. In Arabidopsis, the leptotene stage is identifiable by the initiation of programmed double-strand break formation, the first step of recombination and the beginning of the bouquet stage, where telomeres attach to the nuclear envelope, facilitating the homologous search.
    2. **Zygotene**: During zygotene, homologous chromosomes continue to condense and begin to align with each other. This alignment is followed by synapses formation, facilitated by forming a protein complex called the synaptonemal complex. In Arabidopsis, the synaptonemal complex progressively forms between homologs, initiating at recombination sites and extending along their lengths to ensure accurate chromosomal pairing.
    3. **Pachytene**: The pachytene stage is characterized by the full synapsis of homologous chromosomes. The chromosomes are fully condensed and paired along their entire length, with the synaptonemal complex stabilizing the pairing. Double strand breaks are repaired as non-crossovers or cross-overs, where non-sister chromatids exchange genetic material, forming chiasmata visible as X-shaped structures under a microscope. This genetic exchange is a crucial source of genetic diversity in Arabidopsis thaliana.

4. **Diplotene**: In diplotene, the synaptonemal complex disassembles, and homologous chromosomes begin to separate slightly. However, the homologous chromosomes remain attached at chiasmata at the crossover sites. During this stage, chromosomes start to desynapse, and the chiasmata are more clearly visible.
5. **Diakinesis**: The final stage of Prophase I is diakinesis. Chromosomes are at their maximum condensation level, making them highly visible under the microscope. Chiasmata move toward the ends of the chromosomes in a process called "terminalization." The nuclear membrane begins to break down, and the meiotic spindle forms in preparation for metaphase I. Arabidopsis thaliana cells exhibit precise chiasma distribution, crucial for adequately segregating homologous chromosomes during the following meiotic division.

- **Metaphase I:** Homologous chromosomes align at the metaphase plate with the spindle apparatus firmly attaching to their kinetochores. In Arabidopsis, this alignment is crucial for the equal segregation of chromosomes, which is tightly regulated to prevent aneuploidy.
- **Anaphase I:** During anaphase I, homologous chromosomes are pulled apart to opposite poles of the cell. Arabidopsis meiocytes exhibit the reductional division where homologs, rather than sister chromatids, are segregated - a defining feature of meiosis.
- **Telophase I and Cytokinesis**: Telophase I sees the decondensation of chromosomes and the reformation of the nuclear membrane, often simultaneously followed by cytokinesis, which divides the cell into two.
- **Meiosis II:** After the first meiotic division, meiosis II begins without another round of chromosomal replication. It follows a similar phase progression as mitosis, leading to the segregation of sister chromatids during anaphase II and eventually resulting in four genetically distinct haploid cells after telophase II.

This comprehensive understanding of male meiosis in Arabidopsis thaliana allows us to better analyze our live cell imaging datasets. The intricate details of each meiotic stage will be instrumental in guiding our AI algorithms, allowing us to decipher the dynamic processes occurring within these cells accurately. This depth of knowledge is particularly valuable as we explore the nuances of meiotic progression in our targeted Arabidopsis variants. The specific characteristics of meiosis in Arabidopsis, such as the visible chiasmata in diplotene or the detailed chromosomal alignment in metaphase, present both challenges and opportunities for our imaging techniques. These stages offer critical checkpoints for our AI models to identify and analyze, enhancing the precision of our study. Particular attention will be paid to stages like pachytene and metaphase I, where key genetic exchanges and alignments occur, pivotal in genetic diversity and stability in our Arabidopsis variants.

2.4 Confocal live-cell imaging microscopy

Live cell imaging using confocal microscopy provides invaluable insights into the dynamic biological processes within Arabidopsis thaliana. This method, central to our exploration of meiotic stages, utilizes fluorescent markers to deliver precise, high-resolution observations of cellular activity. The resultant images are not merely visual captures but critical data for subsequent AI analysis.

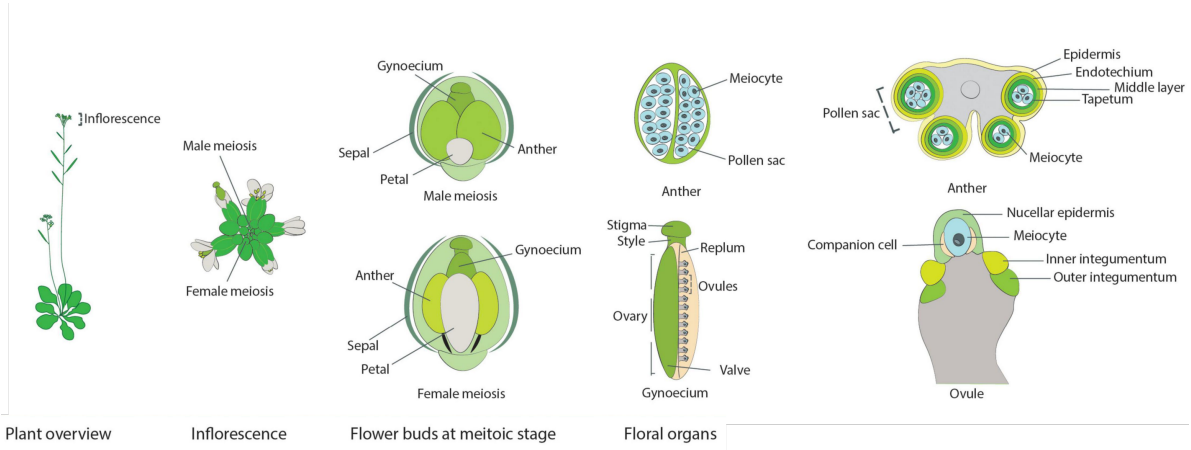

Figure 2.2: Anatomical overview of plant reproductive structures during meiosis *[2]*

The illustration provides a comprehensive view of the plant's reproductive anatomy, showcasing the key structures involved in both male and female meiosis. Starting with an overview of the plant, it details the inflorescence and zooms on the flower buds at the meiotic stage. The diagrams depict the internal organization of anthers and ovules, highlighting the locations of meiocytes within the pollen sacs and ovules, respectively. Further, it presents a cross-sectional view of the anther, showing the layered structure and the positioning of meiocytes, which is critical for understanding the meiotic process in plants.

The precision of confocal microscopy is critical to our study. It achieves detailed visualization by filtering out non-essential light, allowing for focused illumination of specific proteins tagged with GFP or RFP. The parameters of this technique, such as excitation and emission wavelengths, pinhole size, and scan speed, are carefully adjusted to optimize image quality for each specimen.

Our research uses confocal microscopy to examine Arabidopsis's cellular architecture and dynamics. It is especially crucial for observing the behaviour of chromosomes during meiosis, which unfolds within the plant's anthers and ovules. The accompanying image illustrates the various structures of Arabidopsis where meiosis occurs, including the anther and ovule, which house the male and female meiocytes, respectively. These structures are critical for our imaging work, as they are the sites where the complex process of meiosis can be observed.

Fig. 2.2 depicts the distinct components of Arabidopsis, such as the inflorescence, the location of male and female meiosis within the flower structure, and a cross-section of an anther where the meiocytes are located. Understanding the spatial arrangement of these structures is essential for setting up our imaging protocols and interpreting the resulting images. Our confocal microscopy approach is meticulously adapted for prolonged observation to maintain cellular integrity. Time-lapse imaging captures the progression of meiosis over time, while dual-colour fluorescence allows us to simultaneously track multiple cellular components, providing a comprehensive view of the cellular events.

The confocal microscopy setup for studying live cell dynamics in Arabidopsis thaliana is depicted in Fig. 2.3, which details the meticulous preparation and imaging process. In panel A, we see the schematic of a water immersion objective setup, where a thin layer of water and a 2% agarose gel is used to mount the specimen in Arabidopsis culture medium (ACM), ensuring optimal imaging conditions. Panel B displays the preparation of the flower buds, with a close-up view of a bud positioned for imaging and a detailed view of a meiocyte within

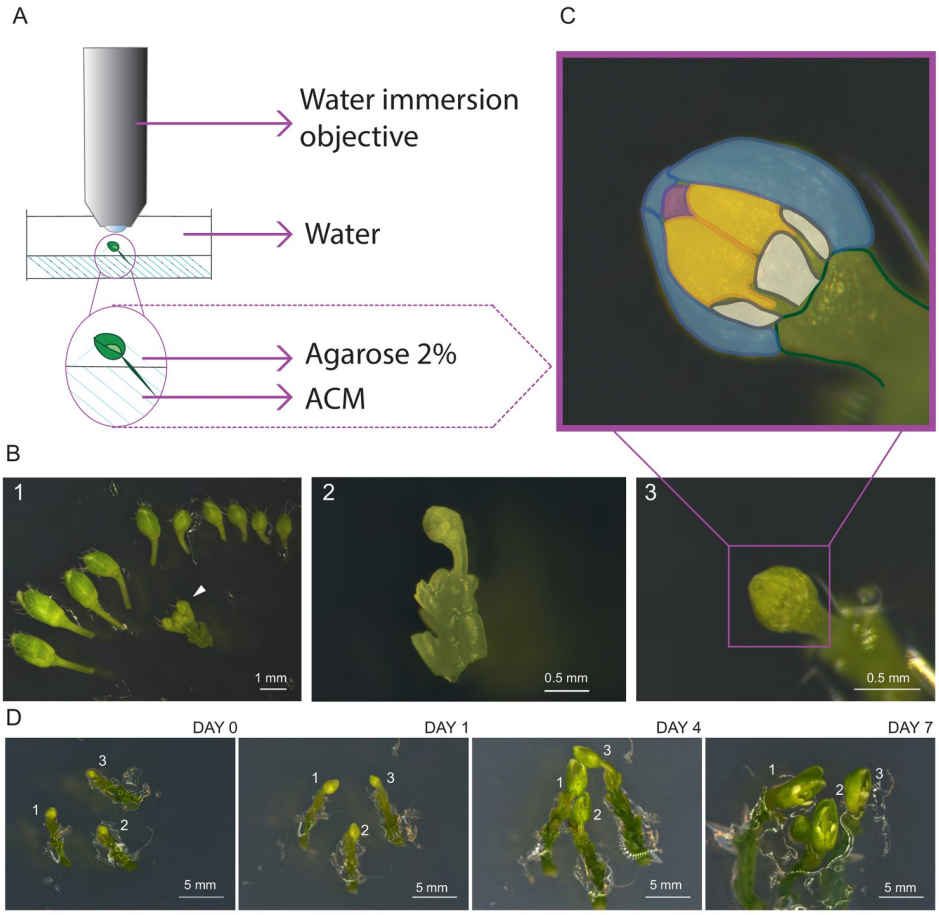

Figure 2.3: Live Cell Imaging Setup and Sample Preparation for Arabidopsis thaliana *[3]*

Panel A illustrates the setup for live cell imaging using a water immersion objective to capture detailed images of the meiocytes within an anther. The schematic shows the anther immersed in agarose gel with Arabidopsis Culture Medium (ACM) to stabilize it during imaging. Panel B displays various stages of flower bud preparation, where 1 indicates a collection of buds ready for imaging, and 2 shows a close-up of a single bud prepared on the imaging slide. Panel C provides a detailed illustration of the cellular components within an anther during live imaging. Panel D demonstrates the growth progression of Arabidopsis thaliana over several days (Day 0 to Day 7), showcasing the developmental changes in the buds and anthers under observation.

it, demonstrating the clarity with which cellular structures can be visualized. An annotated illustration of a bud in panel C delineates various floral parts, clarifying the context of the observed cellular structures. Finally, panel D provides a time-lapse series showing the growth of buds over a week, illustrating the developmental changes captured by our imaging protocol. These visuals guide the imaging process and feed critical data into our AI models, providing the detailed information necessary for an in-depth analysis of meiotic progression.

This imaging technique generates the detailed visual data that fuels our AI algorithms, enabling them to decipher the complex patterns of meiosis. Through these high-resolution images, we gain a deeper understanding of the genetic processes in Arabidopsis thaliana, facilitating a more thorough analysis of plant biology.

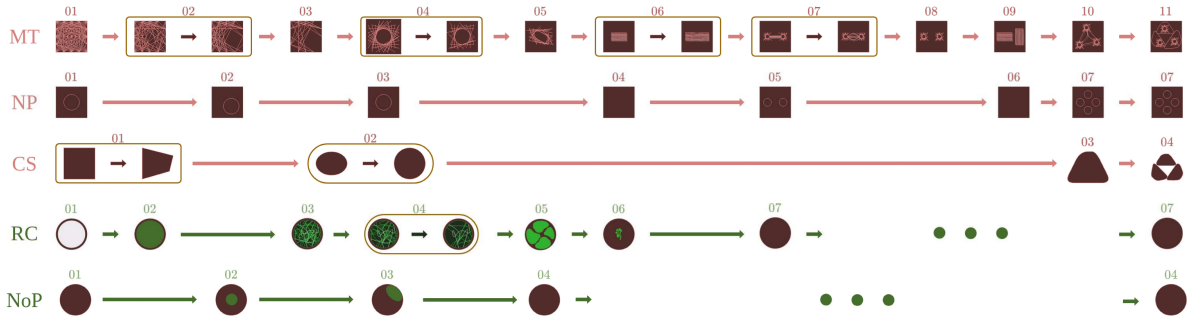

Figure 2.4: *State indicative parameters (sip)*
Cellular features and feature expressions used to characterize (i.e., to classify) the meiocyte states. MT: microtubule cytoskeleton; NP: nucleus position; CS: cell shape; RC: chromatin condensation level; NoP: nucleolus position. The double representation of the penultimate RC and NoP states indicates the state stagnation until the end of the meiosis cycle.

2.5 Defining meiosis phases as state vectors

Transitioning from observing the biological stages to computational analysis, we adopt a quantitative approach by defining meiosis phases as state vectors. This method captures the dynamic cellular changes during meiosis more effectively than traditional fixed-sample analysis. Live cell imaging provides a dataset rich in cellular detail, ideal for computational modelling. Meiotic phases are no longer identified solely by static chromosome configurations but by a series of state vectors representing dynamic cellular conditions.

Traditional phase identification relies on fixed material, observing chromosomal configurations to define stages such as pachytene by fully synapsed chromosomes. However, live cell imaging as applied to *Arabidopsis thaliana* using the KINGBIRD reporter line offers a real-time view of meiosis, where meiocytes are characterized by cell shape, microtubule array, nucleus position, nucleolus position, and chromatin condensation. This live imaging technique, as reported by Prusicki et al. (2019) [3], provides a dynamic representation of cellular states, integrating various cellular components and processes, enabling the differentiation of meiotic states based on several parameters, including cell shape, microtubule array, nucleus and nucleolus positions, and chromatin condensation. The resulting states are not merely static representations but dynamic, integrating the intricate interactions of various cellular processes. This approach provides a refined and detailed view of meiotic phases, particularly cell transitions.

The live imaging data, characterized by distinct cellular parameters, provides a rich dataset for computational modelling. By defining meiotic phases as state vectors, we create a structured framework that AI algorithms can directly interpret. This framework not only captures static chromosome configurations but also the dynamic changes in the cellular environment, which are critical for understanding the process of meiosis.

In meiosis, the progression from one phase to another is marked by distinct changes in chromatin structure, chromosome dynamics, and overall cellular organization. These phases are traditionally determined by observing fixed material and chromosome spreads, where each phase is primarily defined based on chromosome configurations. For example, pachytene, one of the sub-phases of prophase I, is defined by the presence of fully synapsed chromosomes. However, this traditional method of defining meiosis phases has limitations, particularly in capturing the dynamic nature of the process.

This innovative approach allows for the distinct visualization of five key parameters: cell shape, MT array, nucleus position, nucleolus position, and chromosome configurations. The method leverages tubulin to visualize cell shape and MT array, and REC8 reveals chromosome configurations. This analysis identifies multiple states for each parameter, characterized by their specific order and transitions. For instance, cell shape progresses through several distinct forms, culminating in a tetrad of triangular cells. The MT array, showing the most varied states, transitions from a homogeneous distribution to complex structures like half-moon, full-moon, pre-spindle formations, and phragmoplast-like structures during meiotic divisions. Nucleus and nucleolus positions shift in a coordinated manner, indicating changes in the nuclear structure and size, while the nucleolus visibility is linked with REC8 accumulation. Finally, REC8 localization tracks chromatin conformation, aligning with previous immunolocalization studies, and highlights critical meiotic events like chromosome pairing, REC8 removal, and chromosome condensation.

These parameters do not change in isolation but are interconnected, forming a complex matrix that defines meiotic progression. For instance, specific MT arrays, like the full-moon structure, are never found in cells with a rectangular shape, indicating a specific sequence of cellular transformations. This interconnected progression provides a comprehensive view of meiocytes' dynamic and complex changes during meiosis.

Each meiotic datapoint subsequently is encoded as a state vector that includes cell shape, microtubule array, nucleus position, nucleolus position, and chromatin condensation, reflecting the multifaceted nature of the cellular state during meiosis. This encoding is crucial for applying machine learning techniques, which can identify patterns and transitions in these state vectors, predicting the progression of meiosis with high accuracy. Integrating visual classification with fluorescent markers allows for precisely delineating meiotic states. These classified states, interpreted through the dual perspectives of tubulin (RFP) for microtubular structures and REC8 (GFP) for chromosomal dynamics, provide a multidimensional view of meiotic progression. It's this detailed, quantitative description that AI models leverage to learn and predict cellular behaviours during meiosis.

The states of the meiocytes are characterized by visual classification according to the cellular features illustrated in Fig. 2.4. The classified staging furthermore was based on two protein fluorescent markers, tubulin (RFP) and REC8 (GFP); see [3] for details. The RFP marker was used to annotate the microtubular (MT) and the dependent states - the cell shape (CS) and the nucleus position (NP) state. The GFP marker was analyzed to characterize the chromosomal dynamics, i.e., to classify the represented meiocyte according to its RC state and the dependent nucleolus position (NoP) state. Based on the visual assessment of the state for the five features, the overall state of meiocyte k at any timepoint $t \in [0, \ldots, T]$ with T as the final time of progression is represented by the state-indicative vector (siv).

$$siv_{k,t} = [CS_c, NP_n, NoP_o, RC_r, MT_m]\,, \tag{2.1}$$

and, as indicated in Fig. 1, $c \in [1, \ldots, 4]$, $n \in [1, \ldots, 7]$, $o \in [1, \ldots, 4]$, $r \in [1, \ldots, 7]$, and $m \in [1, \ldots, 11]$.

By transforming the visual complexity of meiosis into quantifiable vectors, we lay the groundwork for AI to unravel the patterns of meiotic progression. This quantitative approach does not simplify the biological process but enriches our computational models with the depth and nuance necessary for sophisticated analysis. Ultimately, this confluence of biology and AI opens

new avenues for understanding Arabidopsis thaliana and eukaryotes' cell division fundamentals.

2.6 Live-imaging Dataset

The automated pipeline for analyzing meiosis in *Arabidopsis thaliana* leverages an in-house live-imaging dataset, which is crucial for transitioning from theoretical and biological exploration to computational modelling. This dataset forms the empirical foundation of our AI-driven analysis, providing the necessary diversity to train, validate, and test the robustness of our computational models. Table 2.1 provides an overview of the datasets utilized, highlighting the range of conditions and genetic backgrounds.

The Wild-type dataset-I is central to the thesis, used extensively to train and validate our AI pipeline. In parallel, Wild-type dataset II is vital in evaluating the pipeline's adaptability and generalizability to new, unseen data. The tetraploid and ATM mutant datasets expand the analysis scope, enabling us to examine temporal variations in the meiosis timeline and assess the concordance with established studies in the field. These datasets allow for the application of complex AI algorithms to detect and interpret the intricate cellular events of meiosis, transitioning from static chromosome configurations to dynamic state vectors.

In the following subsection, we detail the primary datasets used to develop the pipeline, while evaluation datasets will be introduced progressively as the thesis narrative unfolds. Dataset-III, captured by a separate imager yet under similar conditions as Dataset-I and -II, provides an additional layer for evaluating the meiotic timeline. The tetraploid dataset allows for examining meiosis temporal variation, and the tetraploid-aTM mutant is studied in the context of these variations.

2.6.1 Wild-type MP dataset

The diploid dataset imaged by Maria Prusicki (MP), featuring annotated wild-type samples, captures images through live-imaging confocal microscopy, steering the selection of parameters crucial for creating an analytical apparatus enhanced by machine learning. This dataset is based on the data introduced by Prusicki et al. [3], which describes the data acquisition process. A total of 35 live cell videos of anther samples were selected to capture the entire meiosis progression, recorded at a frequency that allows for the detailed observation of cellular dynamics.

The anthers in these videos were segmented using threshold-based methods after Gaussian smoothing, refined by median filtering to reduce segmentation artefacts. Within these anthers, 169 *anchored* meiocytes were manually tracked over time, resulting in 10,671 data points representing the position of cells at various time points during meiosis. These data points form the basis for model development, training, and performance evaluation of the different analytical modules in the aMP kit.

2.6.2 Wild-type JDJB dataset

The diploid dataset imaged by Joke de Jaeger Braet (JDJB) consists of 12 videos with 1,923 frames, serving as an independent test set for the aMP pipeline. Captured using the same imaging protocol and conditions as wild type diploid MP dataset, it provides a control for assessing the framework evaluation for meiosis timeline. This dataset was published by Jaeger-Braet et al. [32] and, like Dataset-I, was not used for model development but as a test set to evaluate the pipeline's efficacy. The dataset is not used in individual model performance owing to its manual approach only with microtubule array dynamics.

2.6.3 Tetraploid dataset

The tetraploid dataset, featuring *Arabidopsis thaliana* with four sets of chromosomes, provides a comparative perspective on meiotic progression. The tetraploid-aTM dataset, with a mutation in the ATM gene, offers insights into the impact of genetic variations on meiosis. Including these datasets underscores the model's capacity to handle variations and contributes to an intricate understanding of meiotic timelines for the diploid variants at normal temperatures.

Dataset-type	Imager	Videos	Frames
Wild-type diploid	MP	35	2081
Tetraploid(wild type)	MP	24	2373
Tetraploid(atm)	MvdH & JdJB	6	1928
Wild-type diploid	JdJB	12	1923
Wild-type control	HT	11	1071
TCX	HT	15	2437

Table 2.1: Overview of Arabidopsis thaliana Imaging Datasets:
This table summarizes the different datasets used in the study, detailing the types of *Arabidopsis thaliana* variants, the imager responsible for each dataset, and the quantity of videos and frames within each dataset. The datasets include various wild-type and mutant lines, as well as control and heat-treated samples, providing a comprehensive range for robust AI model training and validation.

2.7 Biological priors

The annotated Wild-type Dataset-I underpins the machine learning apparatus, with visual representations detailed in Fig. 2.4, Fig. 2.5, and Fig. 2.6. A deep understanding of the biological priors that govern this process is essential to develop an AI model that accurately represents the complexity of meiosis. This section outlines these priors, which are critical in guiding the parameterization of our convolutional neural network (CNN) models for the aMP pipeline. These priors inform the selection of AI model parameters and ensure that the model's outputs are biologically meaningful.

- **Anther position and orientation:** Anthers, the floral structures containing the meiocytes, are our first region of interest. The distribution of the abscissa $(x1, x2)$ in Fig. 2.5,

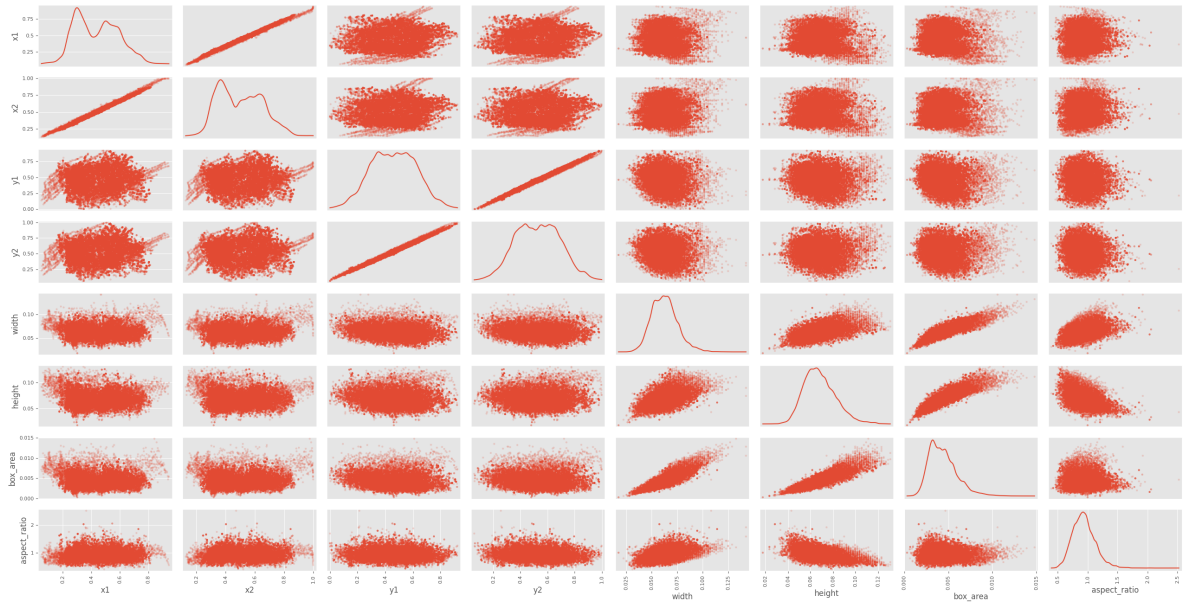

Figure 2.5: ***Dataset property***
The generic properties of the meiocyte in the dataset are described in the diagram. For instance, we can intuitively follow that the cells are divided into 2 anthers from the abscissa plot in the first 2 diagonals. In contrast, the ordinates provided the information that the meiocytes are always centred concerning the video frames.

indicates 2 prominent peaks, ensuring the training dataset consists of 2 anthers in vertical positions. We counter-prove the same, pointing to the distribution of ordinates $(y1, y2)$, which is approximately flat. The information ensures an upper boundary of the detected anthers in an image, excluding false detection. The restricted orientation further influences image augmentation to include other orientations.

- **Meiocyte position and size variation:** Meiocytes, the principal region of interest for our project, provide us with some crucial information and can be found along the diagonal of Fig. 2.5. The abscissa and the ordinates of the bounding box coordinates containing meiocytes show they are present in both the anthers. The bounding box width, height, area and aspect ratio provide different size variations of the meiocyte normalized to the image frame size. The information utilized in setting the aspect ratio and the scales of the meiocyte dimensions facilitates network optimization and filters outliers.

- **Data imbalance:** The distribution of meiocyte data points in the training dataset is left-skewed (Fig. 2.6), with more samples concentrated towards early meiotic stages. The skewed distribution results from a fixed sampling rate used during live cell imaging, which does not align with the inherently variable pace of meiosis progression - slow during prophase-I, extremely short-lived nuclear envelope breakdown (NEB) and fast progression through meiosis-II. Moreover, the manual labelling process inherently involves some classification ambiguity, denoted as the class 00 for each parameter, as annotators encounter challenging cases that defy precise categorization. The data imbalance is alleviated with a weighting factor introduced for each class, and each parameter is appended with the ambiguous class to negate the force staging of a data point.

	box area	aspect ratio	width	height
count	11412			
mean	0.005271	0.974959	0.068627	0.072517
std	0.00351	0.199422	0.018045	0.022848
min	0.000996	0.395062	0.029297	0.02832
25%	0.003437	0.83871	0.057617	0.058594
50%	0.004463	0.963415	0.06543	0.067383
75%	0.005751	1.098765	0.075195	0.080078
max	0.03712	2.396552	0.18125	0.232

Table 2.2: Statistical Analysis of Meiocyte Bounding Box Dimensions
This table comprehensively summarises the bounding box dimensions encompassing meiocytes in the dataset. It includes statistical metrics such as count, mean, standard deviation, minimum, 25th percentile, median (50th percentile), 75th percentile, and maximum for box area, aspect ratio, width, and height. These metrics offer insights into meiocytes' size variability and shape characteristics, vital for precise image analysis and AI model training.

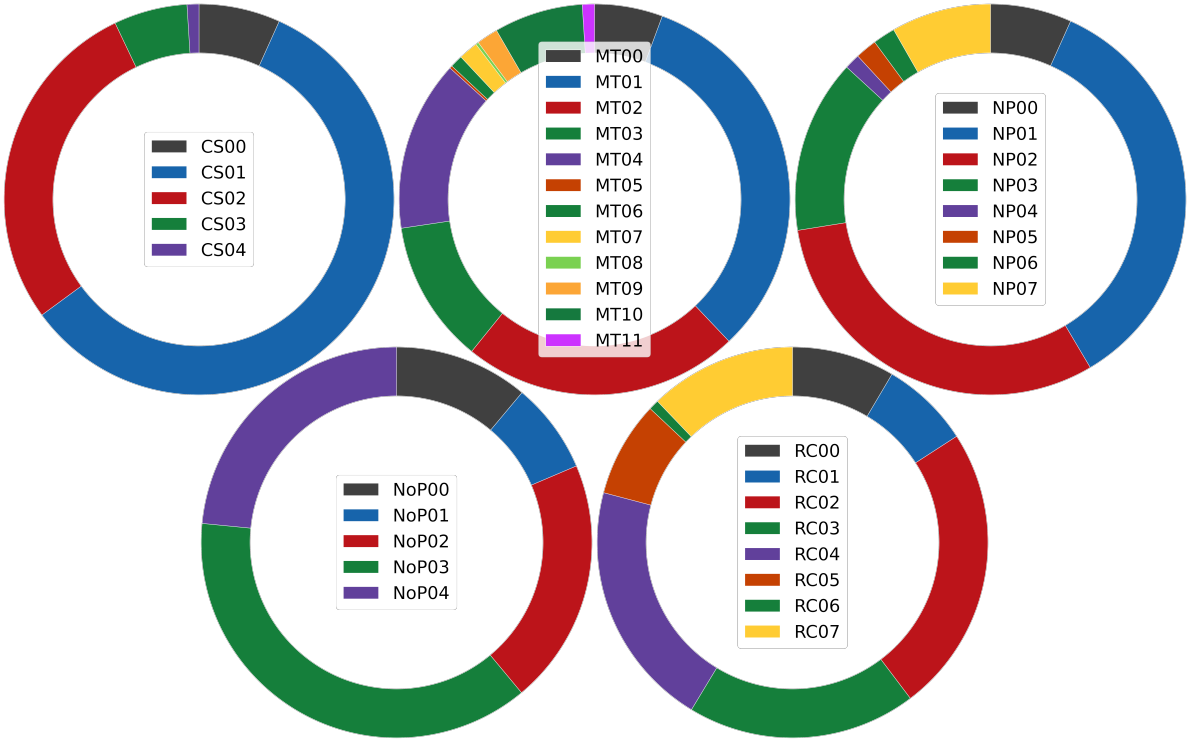

Figure 2.6: ***Parametric meiotic states***
This figure presents the five manually labelled parametric states of meiocytes, with state '00' indicating instances where classification is ambiguous due to indistinct meiocyte characteristics. Variability in these ambiguous classifications arises from the non-characteristic visualization of meiocytes at specific time points. Additionally, there is a notable class imbalance across the labels, attributed to the fixed time intervals used during video acquisition, which do not align with the variable rates of meiosis progression.

2.7.1 Image priors

Understanding the image characteristics is key to fine-tuning the AI model. These priors include:

- **The concept of Anchor Cell:** Live imaging of meiosis necessitates the acquisition of multiple z-stacks, which are sequences of images captured at various focal depths, thus assembling a three-dimensional perspective. This method, however, is subject to the occlusion challenge, wherein adjacent structures or cells may partially or fully conceal aspects of the meiocytes. When processing this data into a 2D format, a critical preprocessing step involves the selection of a single z-stack for each temporal frame. Given the dynamic growth of the flower, there is an inherent shift in the anther's position, implying that consecutive frames might not showcase identical z-stacks. The choice of which z-stack to select hinges on identifying and tracking a set of cells that maintain visual consistency throughout the time series. These precisely tracked cells, selected by the user, are designated anchor cells within the scope of a given live-imaging video. The anchor cells provide the persistent visibility of a meiocyte, forming the upper limit of frames to accurately ascertain when a cell has exited the field of view (fov).
- **Anther movement:** Despite anther immobilizing on a Petri dish during imaging, it often displays significant movement. It falls upon the imaging user to ensure the sample remains within the focal plane of the microscope's field of view. The manual adjustments made by the user and the intrinsic motility of the live anther can result in monotonic and erratic shifts in the sample's positioning. Recognizing these movements is crucial, substantiating the implementation of frame stabilization techniques. Such corrective measures enhance cell tracking accuracy and minimize the confounding effects of sample displacement during live-imaging sessions.
- **Variability in sample illumination:** The illumination intensity during sample imaging is intrinsically linked to the emission characteristics of the fluorescent marker used and the duration for which the sample is exposed to imaging. Specifically, there is an inverse relationship between image intensity and imaging duration. This critical understanding underscores the necessity for adaptive image equalization and contrast stretching techniques. By integrating these adjustments, we can effectively calibrate the image augmentation parameters, thereby standardizing the image intensity across varying conditions and duration of sample imaging. This ensures that variations in fluorescence intensity due to prolonged exposure times are accounted for, maintaining the integrity and consistency of the imaging data.

These biological and image priors are foundational for creating a machine learning-enhanced analytical apparatus. As we transition to developing and applying AI models, these priors ensure that our computational framework is deeply rooted in the biological reality of meiosis in *Arabidopsis thaliana*.

2.8 Data preprocessing

The preprocessing of training samples for the automated Meiosis Progression (aMP) framework is methodically carried out in two primary stages to condition the input data for optimal

performance:

- **Conversion from Multidimensional to Two-Dimensional Imaging:** Confocal microscopy provides raw footage of the anthers at varying focal depths, crucial for the detailed visualization of meiocytes, as necessitated by the biological intricacies previously outlined. The initial step in preprocessing involves manually selecting the most representative z-stack at each temporal point. This curated selection process transforms the multidimensional stack into a coherent two-dimensional image sequence, ensuring that subsequent analyses are based on the most pertinent spatial information.
- **Streamlined Automated Preprocessing:** The next phase of preprocessing employs automated techniques tailored to refine the imaging data:
  - **Frame Resizing:** To unify data dimensions and alleviate computational demands, video frames are resized to a standard resolution of 512×512 pixels. This scaling process incorporates anti-aliasing techniques to filter out high-frequency noise while preserving the integrity of critical image features. Aspect ratio preservation is ensured by padding the shorter dimension before resizing, maintaining the proper morphology of cellular structures.
  - **Contrast Enhancement:** The resized images often suffer from low contrast, which can mask essential details. By executing histogram equalization on the value channel of the images converted to HSV (Hue, Saturation, Value) format, we significantly boost contrast levels. This step is pivotal in normalizing fluorescence intensity variations and enhancing cellular feature visibility. The images are consequently transformed back to the RGB image.
  - **Gray-scaling of RGB Images:** As part of streamlining the image data, the RGB images are converted into gray-scale. This process, crucial in image preprocessing, involves transforming the original three-channel RGB representation (red, green, and blue) into a single-channel grey-scale format. The grey-scale conversion retains essential structural details while making the data easier and more efficient for algorithmic processing, particularly as the fluorophores' coloured annotation is user-specific. To use pre-trained 'imagenet' weights on the grey-scaled image in training our machine learning network, we replicated the grey-scaled image in 3 channels mimicking an RGB image.
  - **Normalization:** The final step in the automated preprocessing sequence involves normalizing the histogram-equalized images. We apply min-max scaling; we adjust pixel values to fall within a [0, 1] range. This normalization is crucial for the consistent treatment of data by CNNs, ensuring that the input features are on a comparable scale for effective pattern recognition and learning.

By applying these data preprocessing steps, we aim to optimize the quality and usability of the microscopic video data for the subsequent analysis. The processed data is now ready for the automation of meiosis progression analysis using CNNs, enabling accurate and efficient study of this critical biological process in Arabidopsis thaliana.

2.9 Data augmentation

We incorporate a strategic data augmentation protocol to bolster the robustness of the automated Meiosis Progression (aMP) pipeline and ensure the convolutional neural network (CNN) models generalize well across varied meiotic imagery. This process increases the volume of training data and introduces a diversity of samples that better represent the range of variability encountered in live imaging of Arabidopsis thaliana meiosis. The augmentation techniques are carefully selected following the biological and imaging constraints previously detailed. The employed augmentation techniques are as follows:

- **Rotational Augmentation**: Considering the radial symmetry of anthers, we apply random rotations to the images. Each frame is rotated by a degree sampled from a uniform distribution within a specified range, ensuring that the CNN learns to recognize meiocytes irrespective of their orientation.
- **Translational Shifts**: Images are translated horizontally and vertically to simulate the natural movement of anthers and account for slight positional displacements. This shift respects the bounds of anther movement observed in the original video sequences to maintain biological integrity.
- **Scaling**: Random scaling of images imitates variations in the z-plane focus, reflecting the differences in meiocyte size due to focal plane adjustments during the imaging process. This scaling is constrained to a range that preserves cellular structures and does not introduce unrealistic proportions.
- **Shearing**: A modest degree of shearing compensates for the potential distortion during the manual adjustment of the anther in the focal plane. This adjustment is subtle to ensure that the cellular morphology remains biologically plausible.
- **Intensity Variations**: Fluctuations in fluorescence intensity due to prolonged imaging times are replicated by adjusting the brightness and contrast of images. This alteration mirrors the inverse proportionality of image intensity to imaging time and the subsequent compensatory preprocessing steps.
- **Elastic Deformation**: Elastic deformations are applied to the images to mimic the biological variability of cell shape and size. This technique introduces a realistic and biologically informed variation, enhancing the model's ability to discern meiocytes under less-than-ideal imaging conditions.
- **Flip Augmentation**: Horizontal and vertical flips represent the natural orientation variability within the anther. This also helps CNN not to be biased towards any specific orientation of meiotic cells.

Each augmentation technique is implemented carefully considering the biological and image priors mentioned above, ensuring the augmented data remains true to the possible physical realities of meiosis in Arabidopsis thaliana and is tabulated in Table. 2.3. Moreover, the magnitude of each transformation is controlled to prevent the introduction of artefacts or biologically implausible features that could mislead the learning process. We significantly enhance the model's exposure to possible imaging scenarios by augmenting our dataset through these varied transformations. This diversity is crucial for developing a resilient and accurate CNN model capable of performing high-fidelity meiosis progression analysis.

Augmentation Technique	Magnitude/Range
Rotational Augmentation	0° to 45°
Translational Shifts	Up to 10% of the minimum image dimension
Scaling	$0.8x$ to $1.2x$
Shearing	Up to 5°
Intensity Variations	Brightness variation by ±20%
Elastic Deformation	Up to 5% deformation
Flip Augmentation	Horizontal and Vertical flips

Table 2.3: Summary of image augmentation techniques and their magnitudes
This table enumerates the various image augmentation techniques applied to the dataset and their corresponding magnitude or range. It includes rotational augmentation, translational shifts, scaling, shearing, intensity variations, elastic deformation, and flip augmentation. These augmentations are critical for ensuring the AI models are exposed to various imaging scenarios, enhancing their ability to generalize and accurately analyze meiotic progression in *Arabidopsis thaliana*.

2.10 Data split strategy

With the preprocessing and data augmentation stages meticulously executed, our dataset is now optimally conditioned for the next phase of our thesis. These steps are crucial in bridging the gap between the biological nuances of meiosis and the computational models that aim to analyze and interpret these intricate processes.

2.10.1 Label stratification

Label stratification is a pivotal machine learning technique, mainly supervised learning. It ensures that each class within a dataset is proportionally represented across various subsets, such as training, validation, and test sets. This approach is essential in handling datasets where class distribution might be imbalanced, preventing any class from being underrepresented or overlooked during the model training and evaluation phases.

Stratification works by maintaining a balanced class distribution, thereby aiding in developing models that generalize well across all classes and avoid biases towards majority classes. This method is crucial in simple classification tasks and plays a significant role in complex scenarios like multi-class and multi-label classifications, where the relationships between classes are intricate.

Proper data splitting techniques, including stratification, are necessary for training classification networks. Representative sampling is crucial to ensure the training set mirrors the real-world scenario the model will encounter post-deployment. Additionally, it's imperative to prevent data leakage between the training and test sets to achieve realistic performance estimations. Uniform application of preprocessing steps across all data subsets is essential to maintain consistency and avoid biases. Ensuring sufficient sizes for validation and test sets is vital for practical model tuning and accurate performance evaluation. Moreover, data randomisation is crucial in eliminating any inherent biases in the dataset.

Description	Count	Meiocyte datapoints	Frames	Anther datapoints
Unique flower videos	19	11412	2081	3623
Train Set	13	9582	1575	2804
Validation Set	3	1255	165	278
Test Set	3	575	341	541

Table 2.4: Distribution of Arabidopsis thaliana Flower Videos and Data Points Across Training, Validation, and Test Sets

This table presents the allocation of unique flower videos and associated meiocyte and anther datapoints among the training, validation, and test sets. The division demonstrates label stratification, ensuring a balanced representation of data across all subsets to facilitate effective training and evaluation of the AI models.

2.10.2 Test-train-validation split

In preparation for developing and evaluating our CNN-based deep learning models, dataset I was partitioned into distinct subsets for training, validation, and testing, allocating a $70\% - 15\% - 15\%$ split ratio. The label stratification strategy ensured the videos' exclusivity (eliminating the overlap in train-validation-test samples) in training the supervised training networks - for classification, segmentation, and object localization tasks. Table. 2.4 shows the dataset used in training the aMP pipeline.

The systematic preprocessing, data augmentation, and strategic dataset division play a crucial role in the subsequent analysis phases of this research. These preparatory steps have been meticulously executed to transform the complex biological phenomenon of meiosis in Arabidopsis thaliana into a dataset suitable for computational analysis. This transformation is pivotal for ensuring that the AI models can accurately interpret the dynamic nature of meiotic progression. The datasets, prepared with precise preprocessing and data augmentation, are crucial for capturing the intricate details of meiotic progression. As we advance, the emphasis shifts to the technical application of image processing methods and AI models designed to compute the meiosis timeline across different Arabidopsis thaliana variants. The forthcoming sections are dedicated to an in-depth examination of the algorithms and computational techniques utilized for dataset analysis to extract significant insights into the meiosis process. This transition signifies a critical juncture where raw data is converted into scientific insights, harnessing advanced AI and machine learning methodologies. These methodologies are integral to our understanding of the varied meiotic timelines in different Arabidopsis strains, addressing the core objectives of our research.

Chapter 3

The aMP pipeline

Single-cell analysis is pivotal for studying individual cells, as it allows for observing cellular heterogeneity and understanding unique cellular characteristics within a population. Such detailed observation is particularly significant for the analysis of meiosis progression in Arabidopsis thaliana. It enables examining each meiocyte's progression through meiosis, an inherently dynamic and individualized process where slight variations in progression and behaviour are expected.

Building on the preprocessed and augmented datasets of Arabidopsis thaliana meiocytes from the preceding chapter, we apply single-cell analysis to elucidate the progression of meiosis. Advanced imaging techniques, like live-cell imaging and fluorescence microscopy, are integrated with neural network-based computational methods to enable frame stabilization, precise cell localization, effective tracking, and accurate classification.

This chapter transitions into a methodological exposition, detailing the neural network architectures and computational strategies employed for dataset analysis. Emphasizing the granular examination of meiotic progression, we aim to extract meaningful insights into the variations of meiotic stages among different cells and to analyze the genetic and molecular bases of these processes at the single-cell level. This comprehensive analysis is instrumental in understanding the complexities of meiotic progression, which is essential for grasping plant fertility's nuances and genetic diversity's generation.

Hence, this chapter marks a significant shift from data preparation to applying sophisticated computational models. Here, we outline the methodologies that transform raw data into scientific insights, underpinning our objective of computing and comparing the meiosis timeline across different Arabidopsis thaliana variants. The methods outlined here are chosen for their efficacy in enhancing analysis and enabling comparative evaluations, the key to our goal of delineating the genetic mechanisms that underlie meiotic variability.

3.1 Deep learning in single-cell analysis

Deep learning techniques have become integral in advancing the field of single-cell analysis, with specific methods being applied for varied purposes. These techniques have been tailored to address the unique challenges posed by single-cell data.

For instance, convolutional neural networks (CNNs) are predominantly used for image-based analyses, as discussed by [33]. These networks excel in processing visual data, making them ideal for tasks like cell segmentation and image classification in cellular microscopy. CNNs have shown remarkable success in extracting meaningful features from intricate biological images, aiding in the quantitative representation of cell characteristics.

Autoencoders and variational autoencoders, highlighted in the studies by [34] and [35], are employed for feature selection and data imputation tasks. These methods effectively reduce dimensionality and identify significant features from the high-dimensional data typical in single-cell studies, thereby denoising the data and imputing missing values, common issues in single-cell datasets.

Generative adversarial networks (GANs) and supervised deep learning models, as surveyed by [35], are used for data augmentation and classification tasks. GANs, in particular, have been instrumental in generating synthetic data for training purposes, thereby addressing the challenge of limited sample sizes in single-cell experiments.

As explored by [36], multimodal deep learning approaches integrate various forms of data in single-cell studies, such as genomics, transcriptomics, and proteomics. These techniques are crucial in multi-omics data integration, enabling a more comprehensive understanding of cellular mechanisms.

The advantages of using deep learning methods in single-cell analysis are manifold. They offer superior performance in handling complex and high-dimensional single-cell data, provide improved feature extraction and classification accuracy, and enable the integration of diverse data types. However, there are also notable drawbacks. Deep learning models often require extensive computational resources and large datasets for training. They can be prone to overfitting, especially when dealing with limited sample sizes. Additionally, many deep learning models, particularly the more complex architectures, act as 'black boxes', offering limited interpretability of their internal workings and decision-making processes.

Deep learning techniques have significantly contributed to advancements in single-cell analysis, each method bringing its strengths to various aspects of the analytical pipeline.

3.2 Our approach

This section delves into the intricacies of the Automated Meiosis Progression (aMP) toolkit, designed explicitly for tracing the intricate microtubular and chromosomal dynamics during the meiosis of *Arabidopsis thaliana*. This discussion extends from the general overview of the aMP toolkit provided earlier, focusing on the rationale behind our methodological choices and the architecture of our deep learning models.

The aMP toolkit is structured into modules addressing distinct live-cell imaging data analysis aspects. From preprocessing confocal microscopy images to the detailed interpretation of meiotic stages, we employ state-of-the-art machine learning algorithms and deep learning models, ensuring precision and reliability in our analyses.

Building upon the preprocessing groundwork laid out in the previous chapter, which enhances and normalizes image quality for better analytical accuracy, the heart of our methodology

revolves around developing sophisticated neural network models. These models are designed explicitly for meiocyte segmentation, localization, and classification tasks. The choice of segmentation mainly, was driven by the need to stabilize anthers and accurately delineate cellular structures, a critical step for reliable downstream analysis. Similarly, object localization-based tracking was selected for its effectiveness in precisely mapping the dynamic movements and transformations of meiocytes throughout meiosis, providing a clearer understanding of cellular interactions and processes.

Classification is treated as a distinct block in our pipeline due to its pivotal role in identifying and categorizing different stages of meiosis, which is fundamental for understanding the progression of these cellular events. By separating these tasks, we ensure dedicated focus and specialized processing for each aspect of the cellular dynamics, leading to a dynamic use of the constituent block of the pipeline, singular or combination of blocks.

A key feature of our methodology is adopting a unified backbone across all neural network models. This decision was made to leverage a common architecture's synergistic learning potential and computational efficiency. Such a unified approach not only streamlines the training process but also ensures consistency in feature extraction and interpretation across different stages of analysis, thereby enhancing the overall robustness and coherence of the toolkit.

This method chapter outlines a detailed and comprehensive blueprint of the aMP toolkit. It serves as both a guide for practitioners in the field of computational biology and a transparent exposition of our research methodology for the broader scientific community. This section bridges the conceptual overview provided earlier with specific technical details, underscoring the rationale behind our methodological choices and the architecture of our neural network models.

3.2.1 Rationale behind the selection of supervised learning models

In developing the Automated Meiosis Progression (aMP) toolkit, prioritising supervised learning models over generative models was a significant methodological decision. This choice was guided by several considerations specific to the nature of the data and the objectives of our study.

Supervised learning models are particularly well-suited for tasks where labelled data is available, and the primary goal is to make predictions or classifications based on this data. In the context of the aMP toolkit, we have access to sufficient labelled data about different stages of meiosis in *Arabidopsis thaliana*. This rich dataset enables the supervised models to learn effectively from the labelled examples, making them ideal for accurately classifying the various stages of meiosis and identifying key cellular features. The direct feedback mechanism inherent in supervised learning, where the model's predictions are continuously compared against actual labels, ensures high precision in classification tasks, which is crucial for the detailed and nuanced analysis required in our study.

On the other hand, generative models, which excel in unsupervised environments where the discovery of underlying patterns and data generation is required, were considered less suitable for our objectives. The primary aim of the aMP toolkit is not to generate new data or explore uncharted patterns without prior labels but to analyze and track known biological phenomena with high precision. Moreover, the complexity and computational intensity of generative models

and their tendency for less interpretable results were seen as potential drawbacks in the context of our specific analytical goals.

Furthermore, using supervised models in the aMP toolkit aligns well with the need for reliable and reproducible results in biological research. The clear and structured learning process of supervised models, governed by labelled data, ensures that they learn exactly what they are intended to learn, reducing the chances of ambiguous or misleading outcomes.

In conclusion, the choice to use supervised models in the aMP toolkit was driven by their compatibility with the nature of our data, the precision required for our analysis, and the overarching goal of delivering transparent, interpretable, and reliable results in the study of meiosis in *Arabidopsis thaliana*. This decision underscores our commitment to employing the most appropriate and practical tools for scientific inquiry.

3.3 Neural network backbone

In deep learning, especially within computer vision, the concept of a "backbone network" plays a pivotal role. These backbone networks are pre-trained convolutional neural networks (CNNs) that are the foundational architecture for various complex tasks. The essence of these networks lies in their capability to extract hierarchical features from input images, where initial layers capture essential elements like edges and textures, and deeper layers discern more intricate patterns. This hierarchical feature extraction is crucial for various applications, from image classification to more advanced tasks like object detection and segmentation.

Backbone networks are mainly instrumental in the practice of transfer learning. This technique involves transferring knowledge gained from one task commonly, a task with an extensive and diverse dataset, such as ImageNet, to a different but related task. Such an approach is especially beneficial when the target task has limited data availability. By leveraging pre-trained models, backbone networks provide a robust starting point, significantly enhancing performance on the new task.

Prominent examples of backbone architectures include VGG, ResNet, Inception, and EfficientNet, each characterized by unique features such as network depth, width, and layer types. These characteristics directly influence the network's performance and computational efficiency. Moreover, these networks offer varying trade-offs between accuracy and computational efficiency, with architectures like MobileNets emphasizing efficiency. In contrast, others like ResNets are designed for deeper and more accurate feature extraction. For meiocyte images, these features are the cell's geometry, microtubular structures, the structure of chromosomes, positional features and other underlying characteristics to delineate different stages of meiosis.

Additionally, backbone networks can be fine-tuned or adapted for specific tasks by modifying the architecture, particularly the final layers, or by altering the training process to focus on features more relevant to the specific application.

In essence, backbone networks in deep learning are cornerstone architectures that underpin the feature extraction process in numerous vision-related tasks. Their use of pre-trained models and adaptability make them invaluable in achieving high performance in a broad spectrum of complex visual recognition scenarios, forming the core of the automated Meiosis Progression (aMP) toolkit. In our work, we use EfficientNet as our common backbone CNN architecture.

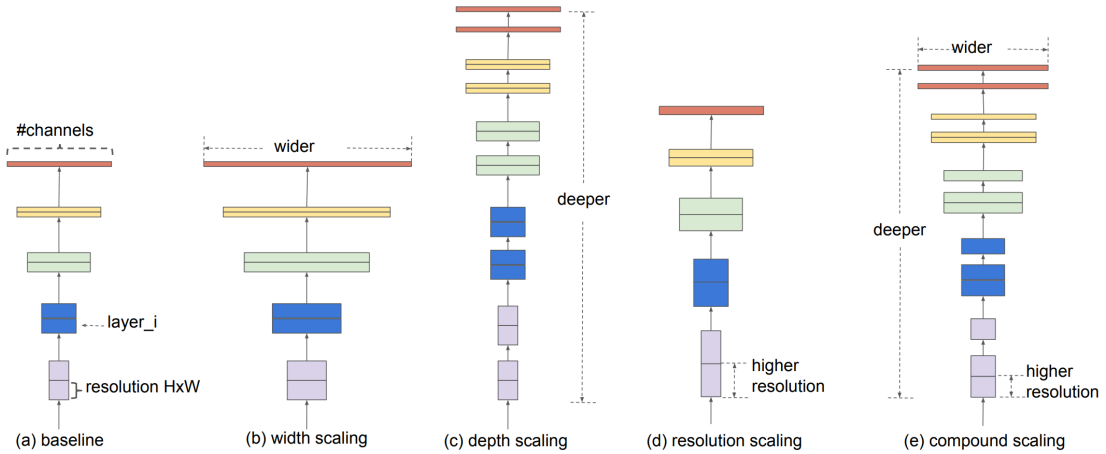

Figure 3.1: ***Model scaling of EfficientNet family.***
a) baseline network, b-d) conventional scaling, e) uniform compound scaling method used in efficientnet with a fixed ratio

3.3.1 Advantages of EfficientNet over other backbone architectures

The introduction of EfficientNet has marked a significant advancement in convolutional neural network (CNN) architectures, addressing key limitations inherent in previous models. In comparison to its predecessors like VGG [37], ResNet [38], Inception [39], MobileNet [40], DenseNet [41], SqueezeNet [42], Xception [43], and YOLO [44], EfficientNet [45] demonstrates several notable advantages.

Firstly, EfficientNet presents a systematic and principled approach to scaling up CNNs, contrasting with the more heuristic methods used in earlier architectures. While architectures like VGG and ResNet primarily focus on increasing the depth of the networks, and Inception introduces a complex combination of convolutions, EfficientNet provides a balanced scaling of all dimensions - depth, width, and resolution (Fig. 3.1. This uniform scaling results in a more efficient and effective network capacity and performance increase.

Secondly, EfficientNet achieves superior accuracy and efficiency, shown in Fig. 3.2. It outperforms other models like MobileNet and DenseNet, which are designed for efficiency, by achieving higher accuracy with a comparable number of parameters. Moreover, compared to SqueezeNet and Xception, which also emphasize parameter efficiency, EfficientNet provides a better trade-off between accuracy and model size.

Furthermore, EfficientDet (built on EfficientNet backbone) effectively overcomes the limitations of model scaling inherent in the YOLO architecture, primarily designed for object detection tasks. While YOLO excels in speed, EfficientNet offers a more versatile architecture that maintains high efficiency without compromising accuracy for various image recognition tasks.

The architectural innovation of EfficientNet lies in its compound scaling method, which simultaneously scales network width, depth, and resolution with a set of fixed scaling coefficients. This approach contrasts with the depth-focused scaling in ResNet or the width and resolution scaling in Inception and Xception, providing a more structured and effective way to improve network performance.

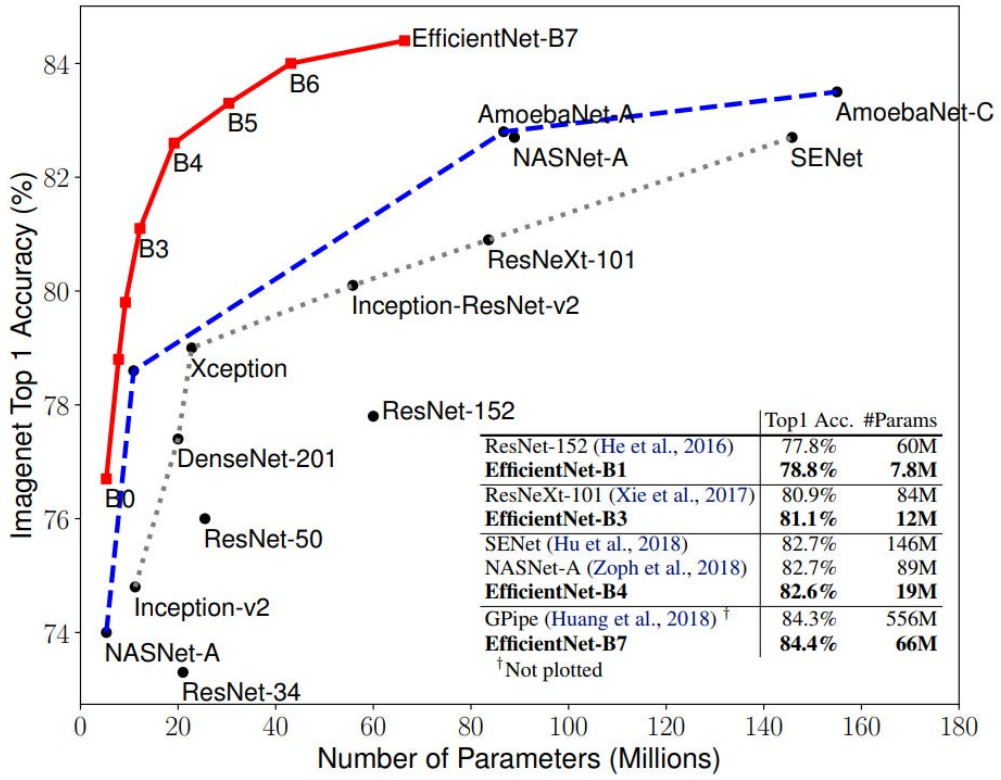

Figure 3.2: ***Comparative accuracy of EfficientNet on Imagenet classification task.***

In summary, EfficientNet represents a significant leap forward in the design of CNN architectures. Its balanced and principled approach to scaling has set new benchmarks in efficiency and accuracy, outperforming its predecessors and offering a versatile solution for a wide range of applications in computer vision.

3.3.2 EfficientNet and ensembling strategy

The EfficientNet architecture comes in several variations, named B0 through B7, each offering a different level of complexity and capacity. The B0 is the baseline network (Fig. 3.3, and each subsequent version, B1 to B7, represents a progressively more complex model that is larger and theoretically more accurate, given more computational resources. One of the distinctive features of EfficientNet is the use of the Swish activation function, which is defined as:

$$\text{Swish}(x) = x \cdot \text{sigmoid}(\beta x), \tag{3.1}$$

where x is the input to the activation function, and β is a parameter that is either learned or set as a constant. This function has been shown to improve the performance of deep neural networks by serving as a smoother alternative to the widely used ReLU activation function.

Another key component of EfficientNet is the MBConv layer, an improved version of the in-

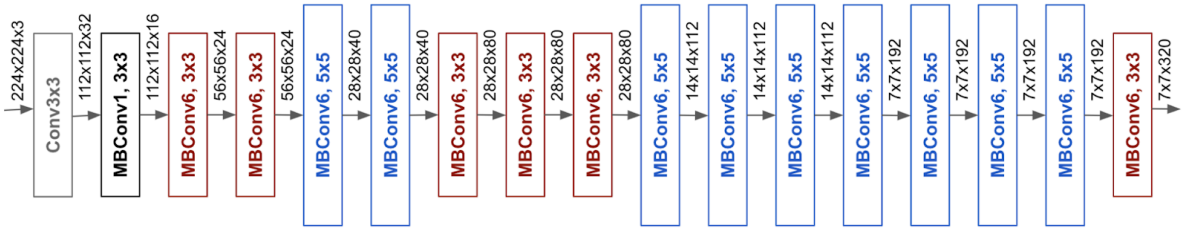

Figure 3.3: ***EfficientNet B0 backbone.***

verted residual structure initially introduced in MobileNetV2. The MBConv layer incorporates lightweight depthwise convolutions to filter features as a form of efficient spatial processing, followed by pointwise convolutions to combine features. The following sequence of operations can mathematically represent the M:

1. Expansion: A 1×1 convolution that expands the input feature map before depthwise convolution.
2. Depthwise Convolution: A 3×3 or larger depthwise separable convolution that acts on each input channel independently.
3. Squeeze and Excitation: An optional block that adaptively recalibrates channel-wise feature responses by explicitly modelling channel interdependencies.
4. Projection: A 1×1 convolution that projects the feature map back to a lower-dimensional space.
5. Skip Connection: An identity connection that is added to the output of the projection layer if the dimensions of the input and output are the same.

Integrating Swish activation and MBConv layers allows efficient combined scaling in EfficientNet, balancing network width, depth, and resolution, leading to better performance than arbitrarily scaling up these factors.

We used 2 more variations - B1 and B2, alongside the baseline B0 model for ensembling learning. The limitation on the availability of computational hardware underlines the limited choice of 3 sub-models. Ensembling is a machine-learning technique that combines multiple models to improve overall performance. The idea is that by combining the predictions from several models, you can capitalize on their strengths and mitigate their weaknesses, forming a comprehensive feature representation and being robust to noise, leading to better generalization and robustness against overfitting.

When it comes to EfficientNet models ensembling B0, B1, and B2 involves using each of these architectures to predict the same input data independently and then aggregating their predictions to form a final output. B0, B1, and B2 are variants of the EfficientNet architecture that differ in depth, width, and resolution, with B0 being the smallest and B2 being larger and more complex. In practice, the predictions from EfficientNet B0, B1, and B2 can be combined using various methods in the different modules of our aMP toolkit.

3.3.3 Backbone feature extractor as plug and play

Exchanging the backbone network within a deep learning toolkit, such as the automated Meiosis Progression (aMP) system, entails replacing the existing feature extraction architecture with an alternative one.

The impetus for such an exchange could stem from several factors, including advancements in neural network designs that offer better performance, the evolving complexity of the dataset, or changes in computational resource availability. For example, one might replace an EfficientNet-B0 backbone with a more advanced EfficientNet variant or a completely different architecture like Vision Transformer to achieve more nuanced feature detection and faster processing times or to fit within a tighter computational budget.

Implementing this change requires careful selection of the new architecture, ensuring it aligns with the specific needs of the meiotic image analysis. Once selected, the new backbone is integrated into the toolkit, often necessitating adjustments in the layers that interface with the rest of the network to accommodate different input or output dimensions. The new backbone may be initialized with weights from a model pre-trained on a relevant dataset to leverage transfer learning for enhanced feature extraction.

The updated model then undergoes retraining, fine-tuning the entire network to tailor the new backbone's capabilities to the specific requirements of meiosis image analysis. Rigorous validation ensures the new architecture maintains or improves the model's performance. This flexible approach to backbone exchange ensures that the aMP toolkit remains adaptable, scalable, and up-to-date with the latest developments in deep learning, thereby maintaining its efficacy in analysing meiotic progression.

3.4 aMP modules

3.4.1 The segmentation module

The segmentation module is the next critical step within the automated Meiosis Progression (aMP) pipeline as we progress from the foundational data preparation elucidated in the previous chapter. This module is pivotal in stabilizing the focal plane across the series of live-cell imaging frames of Arabidopsis thaliana. Ensuring a stable focal plane is imperative for consistent and clear visualization of the meiocytes, which is fundamental for their precise localization and tracking throughout the dynamic stages of meiosis. The following sections will delve into the robust methodologies and approaches that underpin the segmentation module, detailing their contribution to maintaining the stability of the focal plane and enhancing the accuracy of the meiosis timeline analysis.

3.4.1.1 Existing method of analysis

A continuous and systematic methodology is employed to segment anthers using ImageJ/FIJI to ensure accurate and reproducible results. The process begins with pre-processing, where the raw high-resolution images are converted to grayscale to simplify the segmentation process. Fol-

lowing this, a thresholding technique is applied; this can be done through automated algorithms available in ImageJ/FIJI, such as Otsu's method, which calculates an intensity threshold that separates anthers from the background. The binary image resulting from this threshold acts as a preliminary mask, distinguishing the regions of interest (anthers) from the surrounding tissue.

After thresholding, the 'Create Mask' feature solidifies this distinction, producing a mask where the anthers are highlighted. With the binary mask in place, the 'Analyze Particles' function quantifies features of the segmented anthers, allowing for the exclusion of artefacts based on size or shape criteria. This refinement step is essential, particularly when dealing with clusters of anthers or overlapping structures.

The refined binary masks are managed using the ROI Manager, a versatile tool in ImageJ/FIJI that facilitates the editing, measurement, and tracking of multiple-segmented regions across a series of images. To further improve the mask's fidelity, binary operations such as 'Dilate', 'Erode', and 'Close' can be performed, which adjust the mask's boundaries to reflect the anther's shape more accurately.

In some cases, combining multiple masks becomes necessary to segment anthers with varying characteristics within the same image. Logical operators like AND and OR are available in ImageJ/FIJI to merge these masks into a comprehensive representation of the regions of interest. Once the final mask is prepared, it is applied to the original or a stack of images to analyze the anthers exclusively. This masking strategy culminates in the visualization phase, where the mask is overlaid on the original image to verify the segmentation's precision.

3.4.1.2 U^2-Net for anther segmentation

The U^2-Net architecture, introduced by [46], represents a significant advancement in image segmentation, especially in salient object detection. When compared to its U-Net-based predecessors and variants such as the original U-Net [47], V-Net [48], Attention U-Net [49], and U-Net++ [50], U^2-Net exhibits several unique advantages.

While revolutionary in biomedical image segmentation, the seminal U-Net architecture primarily focuses on medical applications and lacks the complexity required for natural image segmentation [47]. Its extensions, like 3D U-Net and V-Net, enhance its capability in three-dimensional data but remain focused on medical imaging [51] [48]. With its attention-guided mechanism, Attention U-Net improves performance in certain image areas but may not consistently excel in salient object detection across diverse scenarios [49]. U-Net++, despite its sophisticated structure, U-Net++ does not fundamentally alter the approach to feature extraction [50].

In contrast, U^2-Net employs a nested U-structure, enabling the extraction of detailed, hierarchical features at multiple scales [46]. This design is particularly effective for salient object detection in complex scenes, where traditional U-Net architectures might falter due to their more uniform feature extraction approach. U^2-Net's architecture is optimized for salient object detection, offering an enhanced ability to discern object boundaries and contextual nuances essential in natural images.

Therefore, while U-Net and its derivatives have significantly contributed to image segmenta-

tion, U^2-Net's distinct nested architecture and focus on salient object detection provide clear advantages for applications in natural image processing, where precise object delineation and complex feature representation are crucial.

3.4.1.3 Training strategy

Arabidopsis thaliana, a model organism in plant biology, features a complex flower structure, including anthers, which are pivotal for studying plant reproduction. An anther typically contains four microsporangia, where male gametophytes develop. Adjacent to these gametophytes lies the tapetum, a nourishing tissue essential for pollen development. In microscopic imaging of these anthers, capturing multiple anthers within a single frame is common, necessitating the precise distinction of individual anther structures. This is particularly crucial when focusing on the tapetum and meiocyte regions, where salient features must be accurately segmented from the complex floral background.

Given the intricate structural details of *Arabidopsis thaliana* anthers and the requirement to preserve spatial information, especially in cases where two anthers are located in one frame, the choice of U^2-Net [46] for saliency detection is highly justified. U^2-Net, with its unique capability to detect and segment salient objects in complex images, is ideal for focusing exclusively on the tapetum and meiocyte regions. Its nested U-structure captures the features necessary for differentiating these critical areas from the surrounding floral tissue.

For the training of U^2-Net, we employed the Adam optimizer [52] for its efficacy in handling sparse gradients, which is common in segmenting high-resolution floral images. The Focal Tversky loss function [53] was utilized to address the class imbalance typically present in these images, where the regions of interest (tapetum and meiocytes) occupy a relatively small portion of the frame. This loss function ensures a heightened focus on these smaller yet crucial regions.

The Continuous Dice coefficient, first introduced in [48], was selected as the accuracy metric, which offered a refined approach for evaluating segmentation models, which is advantageous over the traditional Dice coefficient in several aspects. The standard Dice coefficient is defined as

$$DSC = \frac{2 \times |X \cap Y|}{|X| + |Y|} \tag{3.2}$$

where X and Y represent the binary prediction and ground truth segmentation, respectively, assesses the model's performance in a binary manner. However, the Continuous Dice coefficient, often formulated as

$$CDSC = \frac{2 \times \sum_i p_i g_i}{\sum_i p_i + \sum_i g_i} \tag{3.3}$$

where p_i and g_i denote the predicted probability and ground truth for each pixel i, incorporates the probabilistic nature of pixel classification. This continuous formulation leads to a heightened sensitivity to subtle variations in segmentation, making it particularly useful for models where precision is key. It offers smoother gradients, conducive to the gradient-based optimization methods prevalent in deep learning. This smoothness facilitates more stable and effective training of neural networks, evident in improved convergence behaviours. Moreover, the Continuous Dice coefficient's nuanced evaluation metric is more aligned with soft segmentation tasks, common in medical imaging, where boundaries are not strictly defined. Its ability to distinguish between varying levels of prediction certainty adds depth to the performance assessment, a feature absent in the standard Dice coefficient.

An exponential learning rate decay is implemented to adjust the learning rate during training. This approach gradually reduces the learning rate, allowing the optimizer to make smaller, more precise updates to the model weights as training progresses. Such a strategy helps stabilise the training process in its later stages, leading to better convergence.

We implement an early stopping mechanism to prevent overfitting and ensure efficient training. Training is terminated based on the validation loss value - if the validation loss does not improve for a pre-defined number of epochs, the training process is stopped. This criterion ensures that the model training ceases when it begins to overfit, thereby preserving the model's generalizability.

3.4.1.4 Anther segmentation

Given a preprocessed video denoted as $\mathbb{V}$ with its constituent frames represented by $\mathbb{V}_{t=0,\dots,T}$, the segmentation block's primary objectives are twofold. Firstly, it aims to delineate an appropriate mask $\mathbb{M}_{t=0,\dots,T}$ that encapsulates the anther area pertinent for localizing meiocytes. Secondly, it seeks to rectify any positional shifts of the anther across the video sequence, thus facilitating the subsequent tracking of meiocytes. Anther masking facilitates focused analysis on areas of interest, excluding irrelevant background information.

3.4.1.5 Anther Motion Compensation

After segmentation, we address the anther's motion using the segmented masks $\mathbb{M}_{t=0,\dots,T}$. To stabilize the anther's position across frames, we compute the centroid $C(\mathbb{M}_t)$ for each frame and estimate the displacement vector ΔC_t relative to the first frame:

$$\Delta C_t = C(\mathbb{M}_t) - C(\mathbb{M}_0) \tag{3.4}$$

To compensate for the anther's motion, we define a transformation τ that combines both translational and rotational adjustments. Specifically, τ is a composite of translation $\vec{T}_t$ and rotation R_t based on ΔC_t.

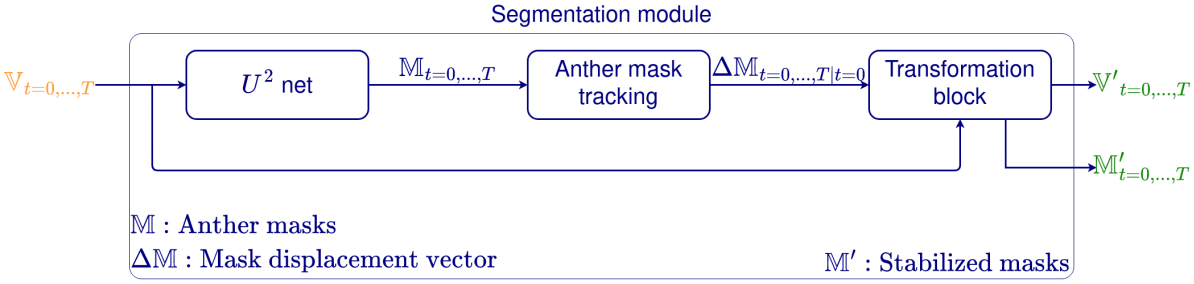

Figure 3.4: ***Structure of the segmentation module.***
Within the input video $\mathbb{V}_{t=0,\dots,T}$, the anther regions that are relevant to meiocyte localization are segmented. Based on the segmented areas $\mathbb{M}_{t=0,\dots,T}$, the anther motion is estimated and compensated, resulting in motion-compensated videos $\mathbb{V}'_{t=0,\dots,T}$ and $\mathbb{M}'_{t=0,\dots,T}$.

For frame t, the translational transformation $\vec{T_t}$ and the rotational transformation R_t are defined as follows:

$$\vec{T_t} = \begin{bmatrix} 1 & 0 & \Delta x_t \\ 0 & 1 & \Delta y_t \\ 0 & 0 & 1 \end{bmatrix} \tag{3.5}$$

$$R_t = \begin{bmatrix} \cos(\theta_t) & -\sin(\theta_t) & 0 \\ \sin(\theta_t) & \cos(\theta_t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.6}$$

where Δx_t and Δy_t are the x and y components of ΔC_t, and θ_t is the angle of rotation.

The composite transformation τ_t applied to each frame is the product of $\vec{T_t}$ and R_t:

$$\tau_t = \vec{T_t} \cdot R_t \tag{3.7}$$

The inverse transformation τ_t^{-1} is then applied to each frame and its corresponding mask to stabilize the anther's motion:

$$\mathbb{V}'_t = \tau_t^{-1}(\mathbb{V}_t) \tag{3.8}$$

$$\mathbb{M}'_t = \tau_t^{-1}(\mathbb{M}_t) \tag{3.9}$$

This results in a stabilized sequence of images $\mathbb{V}'_{t=0,\ldots,T}$ and masks $\mathbb{M}'_{t=0,\ldots,T}$, where the anther maintains a consistent position and orientation relative to the first frame.

3.4.2 Quantitative evaluation of motion stabilization

Displacement Calculation: We calculate the displacement of anthers in each frame after applying the inverse transformation τ^{-1}. The displacement D_t for frame t is computed as the Euclidean distance between the transformed centroid position and the original centroid position in the first frame:

$$D_t = \sqrt{(\Delta x'_t)^2 + (\Delta y'_t)^2} \tag{3.10}$$

where $\Delta x'_t$ and $\Delta y'_t$ are the x and y components of the centroid displacement after applying τ_t^{-1}.

Assessment of Stabilization Effectiveness: The stabilization effectiveness is determined by comparing the displacements before and after applying the motion compensation. We calculate the average and standard deviation of D_t across all frames and categorize the results into different classes based on the extent of stabilization achieved.

Additional Evaluation Metrics: To further assess the stabilization process, we employ traditional image segmentation metrics:

- *Intersection over Union (IoU):* Measures the overlap between the segmented anther regions in the stabilized and the first frames.
- *Precision:* Evaluates the accuracy of the segmented anther regions in the stabilized frames.

- *Recall:* Assesses the completeness of the anther region segmentation in the stabilized frames.
- *F1 Score:* Provides a balance between precision and recall, offering a single metric that encapsulates both aspects.

The combined use of displacement analysis and traditional metrics like IoU, Precision, Recall, and F1 Score offers a comprehensive evaluation of the stabilization process. This multifaceted approach ensures the reliability and accuracy of the stabilization, crucial for the subsequent analysis of meiotic progression in *Arabidopsis thaliana*.

Having established a robust framework for segmenting and stabilizing anther regions, our pipeline now transitions to the critical task of meiocyte localization. Anthers' segmentation and motion compensation provide a stabilized and refined context, crucial for the precise detection and localization of meiocytes, ensuring that the localization module operates on high-fidelity data, where the meiotic cells are preserved and made distinctly identifiable. Within this refined visual landscape, established through advanced segmentation techniques, our meiocyte localization module leverages the power of EfficientDet models to pinpoint the positions of these key cellular entities accurately. This continuity from segmentation to localization is not just a sequential progression but a testament to our pipeline's integrated and modular approach in dissecting the complex phenomenon of meiosis in Arabidopsis thaliana.

3.4.3 The meiocyte localization module

The accurate localization of meiocytes within the stabilized video frames $\mathbb{V}'_{t=0,\dots,T}$ is a critical precursor to effective tracking. This process is facilitated by a single-shot detection framework that leverages the motion-compensated frames as input.

3.4.3.1 Advantages of object localization over-segmentation in cell imaging

Object localization is over-segmentation for locating cells in image frames due to several key advantages. Primarily, it is computationally more efficient, as segmentation necessitates pixel-level classification, which is resource-intensive, especially for high-resolution images. In contrast, object localization, which typically involves identifying bounding boxes or key points, is less computationally demanding. Furthermore, the annotation and training process for object localization is generally simpler and faster than the detailed, pixel-wise annotations required for segmentation. This simplicity translates to quicker model training and less manual labour in preparing datasets.

Another significant advantage of object localization lies in its sufficiency for specific analytical goals. The precise boundary details provided by segmentation are often unnecessary for cell counting or identifying specific cell types. Object localization provides adequate information by pinpointing the location and quantity of cells, which suffices for our further analyses. This method is also advantageous in scenarios with overlapping or clustered cells, where segmentation algorithms might struggle to delineate individual cells accurately. Object localization models, focusing on identifying central points or bounding boxes, are less affected by such overlapping structures.

Moreover, the speed of object localization models is particularly beneficial for real-time analysis, such as in live cell imaging. This rapid processing capability is essential in dynamic environments where cell behaviours continuously evolve. Lastly, object localization models are adaptable to various magnifications and imaging conditions, maintaining their accuracy even when the scale of the image changes. This versatility makes them ideal for a wide range of cell imaging scenarios, where detailed boundary delineation is not critical and computational efficiency or speed is a priority.

3.4.3.2 Choice of EfficientDet

EfficientDet [54] stands out amongst object localization networks for its distinctive features and performance, especially when compared to networks like YOLO [44], SSD [55], RCNN [56], Faster RCNN [57], and CenterNet [58]. One of the critical advantages of EfficientDet is its efficient scaling of model complexity, which is achieved through a compound scaling method. This method uniformly scales the network's depth, width, and resolution, a feature not present in networks like YOLO or SSD, which tend to scale these dimensions heuristically.

Moreover, EfficientDet introduces the Bidirectional Feature Pyramid Network (BiFPN), which allows easy and fast multi-scale feature fusion, enhancing feature learning capabilities. This aspect of EfficientDet is particularly advantageous over networks like RCNN and its variants, where feature pyramid networks are either absent or unidirectional. The BiFPN enables more effective feature-level interactions than traditional FPNs used in Faster RCNN, leading to improved performance in object detection tasks.

EfficientDet also exhibits better parameter efficiency and higher accuracy than the YOLO series, which, although known for their speed, do not scale as efficiently in terms of model size and accuracy. In contrast to SSD, EfficientDet achieves a better trade-off between speed and accuracy. Due to its more advanced feature fusion strategy, the network outperforms SSD in detecting objects at multiple scales.

The architecture of EfficientDet, particularly with the inclusion of BiFPN, demonstrates superior performance in handling objects of various sizes, a common challenge in object detection tasks. This is a significant improvement over CenterNet, which, while efficient in detecting objects as key points, might not capture scale variations as effectively as EfficientDet.

EfficientDet's balanced scaling approach, combined with the innovative BiFPN and its ability to handle multi-scale feature fusion more effectively than other networks, underlines our work's choice of neural network.

3.4.3.3 Training methodology

In our study, we implemented the training of EfficientDet models D0, D1, and D2, each with varying levels of complexity and capacity, tailored to specific object detection needs. The training process was carefully designed, incorporating advanced optimization techniques and loss functions to enhance the models' performance and accuracy.

Anchor scales and ratios play a pivotal role in object detection models by defining a set of reference boxes, or "anchors," which the model uses to predict the presence and location of

objects. The scales determine the size of these anchors, while the ratios define their aspect. For meiocyte detection, where the cells may vary in size due to different stages of development or varying imaging conditions, selecting appropriate scales is essential to ensure the model can detect smaller and larger cells. Similarly, the aspect ratios of the anchors must align with the typical shapes of meiocytes, which can be elongated or circular, depending on the specific developmental stage. The anchor scales and ratios were chosen in our methodology based on the specifics provided in the predefined dataset description section, ensuring that the models are well-attuned to the characteristics of the target objects.

The Adam optimizer [52] was selected for its efficiency and effectiveness in handling sparse gradients, a common scenario in object detection tasks. For the loss functions, we employed the Focal Tversky loss [53] for objectness score optimization and the adaptive smooth L1 loss for bounding box regression. The Smooth L1 loss offers several advantages over the traditional L1 loss, particularly in handling outliers and gradient stability. Mathematically, the L1 loss is defined as:

$$L1(x) = |x|, \tag{3.11}$$

where x is the difference between the predicted value and the ground truth. While straightforward, L1 loss is sensitive to outliers as it linearly increases with the error, leading to potentially unstable gradients during training.

In contrast, the Smooth L1 loss is formulated as follows:

$$\text{Smooth}_{L1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1, \\ |x| - 0.5 & \text{otherwise.} \end{cases} \tag{3.12}$$

This hybrid loss function behaves like an L1 loss for more significant errors (when $|x| >= 1$) but transitions to a squared (L2) loss for more minor errors, reducing the impact of outliers on the training process. The quadratic region near zero helps mitigate the instability in gradients, especially for minor errors, by avoiding the abrupt change in gradient values that L1 loss exhibits. This makes the Smooth L1 loss particularly suitable for regression problems in deep learning, as it combines the robustness of L1 loss with the stability of L2 loss, leading to more effective and stable training [59]. In object detection tasks, where bounding box regression is critical, the Smooth L1 loss ensures that the model is less sensitive to inaccuracies in box coordinates, providing a balance between precision and robustness against noisy data.

The mean Average Precision (mAP) was utilized as the accuracy metric, providing a comprehensive measure of the model's precision and recall across different object classes and scales. This metric is particularly relevant in object detection,

Mean Average Precision (mAP) [60] evaluates object detection models' performance. The metric effectively encapsulates precision and recall, two critical aspects of object detection where the goal is to identify meiocytes and precisely localize them correctly. Precision measures the accuracy of the detections (the proportion of true positives among all detected objects), while recall assesses the model's ability to detect all relevant objects in the image.

Mathematically, Average Precision (AP) for a single class is calculated as the area under the

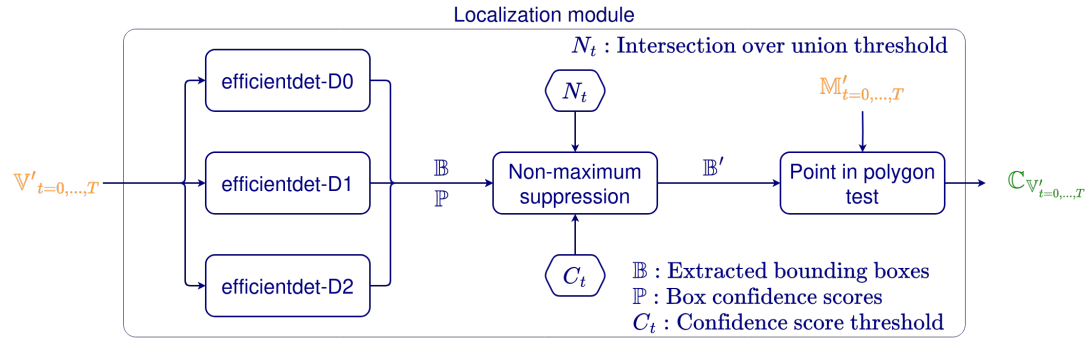

Figure 3.5: ***Structure of the meiocyte localization module.***
Meiocyte candidates are identified using an ensemble of EfficientDet networks. The localized meiocyte candidates are double-checked according to their position, which should be within the anther regions identified in the segmentation module.

precision-recall curve, generated by plotting precision values against recall values at different thresholds. The mean Average Precision (mAP) is then the mean of APs calculated for all classes:

$$mAP = \frac{1}{T}\sum_{t=1}^{T} AP_{\mathrm{IoU}_t} \tag{3.13}$$

where T is the total number of IoU thresholds considered, and AP_{IoU_t} is the AP calculated at the t-th IoU threshold. In the COCO evaluation, a common practice is to average the APs calculated at different IoU thresholds ranging from 0.5 to 0.95 at an interval of 0.05.

3.4.3.4 Inferring the meiocyte locations

The pipeline initiates with the input of confocal microscopic images $V_{t=0,\dots,T}$, spanning from the initial time point $t = 0$ to the final time point $t = T$. For detecting meiocytes, we utilize a series of EfficientDet models, D0, D1, and D2, each providing a unique depth of complexity and scale adaptation, ensuring robust detection across various meiocyte sizes and morphological features.

Non-Maximum suppression technique

Non-maximum Suppression (NMS) is an essential ensembling technique in object detection, employed to refine the outputs of multiple EfficientDet model predicting the potential meiocyte locations, outputting a set of bounding boxes B along with associated confidence scores P

The process begins with an object filtering stage where proposed bounding boxes with confidence scores falling below a predefined threshold, CT, are discarded. This step reduces the set of potential bounding boxes to B'', and the corresponding confidence scores to P'', streamlining the subsequent ensembling process.

NMS then proceeds to ensemble the detections by examining the filtered bounding boxes.

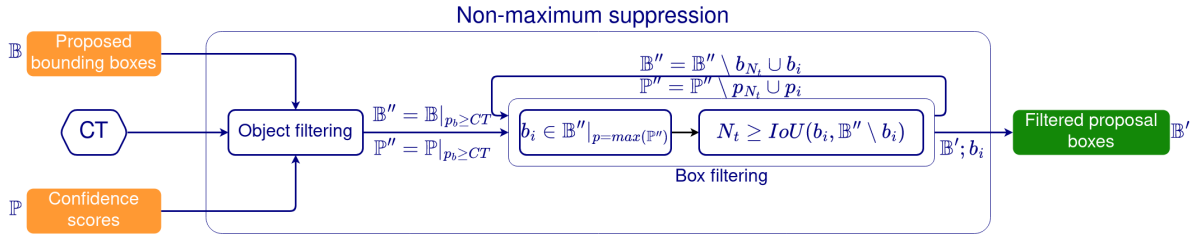

Figure 3.6: ***Structure of the non-maximum suppression block.***
The bounding boxes and the corresponding confidence scores are passed through object filtering to filter out boxes with low confidence; after that, the remaining boxes are filtered out with a high overlap ratio to generate the resulting selection of bounding boxes.

Starting with the box, b_i, that has the highest confidence score, NMS suppresses all other bounding boxes, b_j, that significantly overlap with b_i, as determined by the Intersection over Union (IoU) metric. A bounding box b_j is suppressed if its IoU with b_i exceeds a threshold N_t, thus ensuring that the single most confident bounding box represents each detected meiocyte.

The outcome of applying NMS is a set of filtered proposal boxes, B', constituting the final, non-redundant detections. These boxes are then utilized to localize the meiocytes within the images accurately. Through NMS, the process effectively combines the outputs of different detection models, mitigating the issue of multiple detections for the same object and thus significantly enhancing the accuracy and reliability of the meiocyte detection process.

The NMS process can be formulated as follows:

1. **Aggregation:** Aggregate all bounding boxes and their confidence scores from the three networks into a combined set:

$$B = B^{D0} \cup B^{D1} \cup B^{D2}$$

$$P = P^{D0} \cup P^{D1} \cup P^{D2}$$

2. **Filtering by Confidence Threshold:** Discard bounding boxes with confidence scores below a predefined threshold C_t:

$$B' = \{b \in B \mid p_b > C_t\}$$

$$P' = \{p_b \mid b \in B', p_b > C_t\}$$

3. **Applying Non-Maximum Suppression:** For each bounding box b_i in B':

   - Compute the Intersection over Union (IoU) with all other boxes b_j in B':

$$\text{IoU}(b_i, b_j) = \frac{\text{area}(b_i \cap b_j)}{\text{area}(b_i \cup b_j)}$$

   - If $\text{IoU}(b_i, b_j)$ exceeds a threshold N_t, and $p_{b_i} \geq p_{b_j}$, suppress b_j by removing it from B'.

4. **Final Output:** The remaining bounding boxes in B' after applying NMS represent the localized meiocytes.

This NMS process effectively combines the predictions from the three networks. It refines them to produce a non-redundant set of bounding boxes, improving the accuracy and reliability of the meiocyte localization process.

3.4.3.5 Point-in-Polygon test and sorting with Anther Orientation

After the meiocyte locations are identified through object detection models, the next steps involve verifying these locations within the anther regions and sorting them based on anther orientation. The Point-in-Polygon (PiP) test ensures that each localized meiocyte lies within the predefined anther region. Given a set of bounding boxes B' from the object detection models, each box $b \in B'$ is checked to see if its centroid $C(b)$ lies within the anther mask $\mathbb{M}'_t$ obtained from the segmentation module. Mathematically, the PiP test for a bounding box b in frame t is defined as:

$$\text{PiP}(C(b), \mathbb{M}'_t) = \begin{cases} 1, & \text{if } C(b) \text{ is inside } \mathbb{M}'_t \\ 0, & \text{otherwise} \end{cases}$$

Before sorting the bounding boxes, the predominant orientation of the anthers in the images is assessed. This is determined by analyzing the spatial distribution of anther regions across multiple frames.

Depending on the determined anther orientation, the meiocyte bounding boxes are sorted either in a top-down or left-right manner. Let orientation represent the predominant anther orientation, either 'vertical' or 'horizontal'. The sorted bounding boxes SortedBoxes_t for frame t are then:

$$\text{SortedBoxes}_t = \text{Sort}(B'_t, \text{orientation}) \tag{3.14}$$

, where each contour extracted from the mask can be approximated by a bounding rectangle. The orientation of this rectangle gives us the orientation of the contour. The bounding rectangle is defined by its width w, height h, and the angle θ it makes with the horizontal axis. For a given contour C, the bounding rectangle is calculated as follows:

$$w(C) = \text{Width of the bounding rectangle of } C, \tag{3.15}$$
$$h(C) = \text{Height of the bounding rectangle of } C, \tag{3.16}$$
$$\theta(C) = \text{Angle of the bounding rectangle of } C \text{ with the horizontal axis.} \tag{3.17}$$

and, the orientation of the contour is determined based on the aspect ratio of the bounding rectangle and the angle θ. The contour is considered:

$$\text{Orientation}(C) = \begin{cases} \text{'horizontal'}, & \text{if } w(C) > h(C) \text{ and } (|\theta(C)| \leq 45^\circ \text{ or } |\theta(C)| > 135^\circ), \\ \text{'vertical'}, & \text{if } h(C) > w(C) \text{ and } (45^\circ < |\theta(C)| \leq 135^\circ). \end{cases} \tag{3.18}$$

Finally, the combined process of PiP testing and sorting is implemented for each frame t as

follows:

$$B_t'' = \{b \in B_t' \,|\, \mathrm{PiP}(C(b), \mathbb{M}_t') = 1\}, \tag{3.19}$$

$$\mathrm{SortedBoxes}_t = \mathrm{Sort}(B_t'', \mathrm{orientation}). \tag{3.20}$$

The localization module of our pipeline exports a systematic arrangement of meiocyte bounding boxes $C_{V_{t=0,\ldots,T}}$, each accurately positioned within the segmented anther regions across time-lapse sequences. This process involves verifying the centroid of each bounding box against the anther polygon, ensuring the meiocytes are localized within the relevant biological structures. By employing the advanced detection capabilities of EfficientDet models and application-specific refinement steps, the module effectively catalogues meiocytes, represented by bounding boxes and the corresponding cropped image segments C_t for each video frame $\mathbb{V}_t'$, $t \in [0, \ldots, T]$. This output is pivotal for subsequent biological analyses and tracking the dynamic progression of meiosis in *Arabidopsis thaliana*. Notably, the count of localized meiocytes may vary across frames, reflecting the inherent biological variability in meiotic progression.

3.4.4 The meiocyte tracking module

Building on the accurate localization of meiocytes achieved in the previous module, our focus now shifts to the tracking module. This module aims to meticulously follow these cells through their dynamic developmental stages within the stabilized video sequence $\mathbb{V}'$. The challenge here lies in the morphological changes meiocytes undergo during meiosis in *Arabidopsis thaliana*. Our methodology, detailed in this section, is dedicated to maintaining a consistent identity of meiocytes across successive frames. It is crucial for capturing their morphological transformations and understanding their progression through meiosis. Furthermore, the tracking module is intricately linked to observing and analyzing meiotic stages. By monitoring the trajectory of each meiocyte, we can map out key developmental milestones and identify any deviations from typical meiotic progression.

3.4.4.1 Previous methods to track cells

In a cell tracking challenge [61], which focused on cell segmentation, a fundamental step in cell tracking that delineates cell boundaries or regions within images and, the *segmentation and tracking task*, integrates segmentation with subsequent linking steps, crucial for tracking cells over time. This category includes several methods:

- **Segmentation → Linking (SegLnk)**, where cells are first segmented and linked across frames.
- **Segmentation Linking (Seg&Lnk)**, conducts segmentation and linking simultaneously, providing a more integrated approach.
- **Detection → Segmentation → Linking (DetSegLnk)**, involves initial detection of cellular objects, followed by segmentation and then linking, offering a comprehensive three-step approach to cell tracking.
- **Detection → Linking → Segmentation (DetLnkSeg)**, reverses the order of the last two steps, proposing an alternative workflow.

Segmentation Techniques involve thresholding (S1), region growing (S2), machine learning (S3), and energy minimization (S4), each offering a unique mechanism for accurately delineating cells in images.

Linking Techniques are crucial for tracking cells across sequences and include label propagation (L1), nearest neighbour (L2), graph-based optimization (L3), contour evolution (L4), and machine learning (L5). These techniques are pivotal in establishing continuity and tracking the trajectory of cells over time.

Other elements included in the techniques but not classified under Detection, Segmentation, or Linking are Intensity, Boundary, Spatial statistics, Spatiotemporal statistics, Distance, Overlap, Motion analysis, Shortest path, Minimum cost flow, Probability, Multiple hypothesis, Decision tree, U-Net variant, R-CNN variant, HRNet variant, Siamese tracker, and Graph neural network.

3.4.4.2 Affirming our design choices

This section delves into the approach adopted in our study for tracking cellular dynamics, contrasting it with the conventional methods delineated in the extant literature. This juxtaposition elucidates the advanced nature and potential efficacy of the methodology proposed.

Enhanced detection through EfficientDet

Conventional detection techniques referenced in the literature include thresholding (D1), peak localization (D2), and broad machine learning strategies (D3). While foundational, these methods lack specificity in complex cellular environments. In stark contrast, our study integrates the EfficientDet model for detection. Renowned for its computational efficiency and accuracy, EfficientDet represents a significant advancement over generic techniques, promising enhanced precision and reliability in cell detection as elaborated in 3.4.3.2. In tackling the tracking of meiocytes, we encountered challenges related to the volume of data and the dynamic nature of meiotic cells. Our chosen methodologies and computational strategies effectively address these issues, ensuring accurate and reliable tracking even in complex scenarios.

Novel linking mechanism

Traditional linking techniques, ranging from label propagation to machine learning-based approaches, are effective in general applications but may fall short in meiocyte tracking from live-cell imaging, owing to its structural dynamics. Our research introduces a novel linking mechanism comprising a covariance distance matrix, a search space-restricted overlap measure, and template matching. This methodological triad offers several advantages:

- The **covariance distance matrix** provides an understanding of spatial relationships between cells.
- The **restricted overlap measure** enhances precision in tracking, particularly in cases of overlapping or closely situated cells.

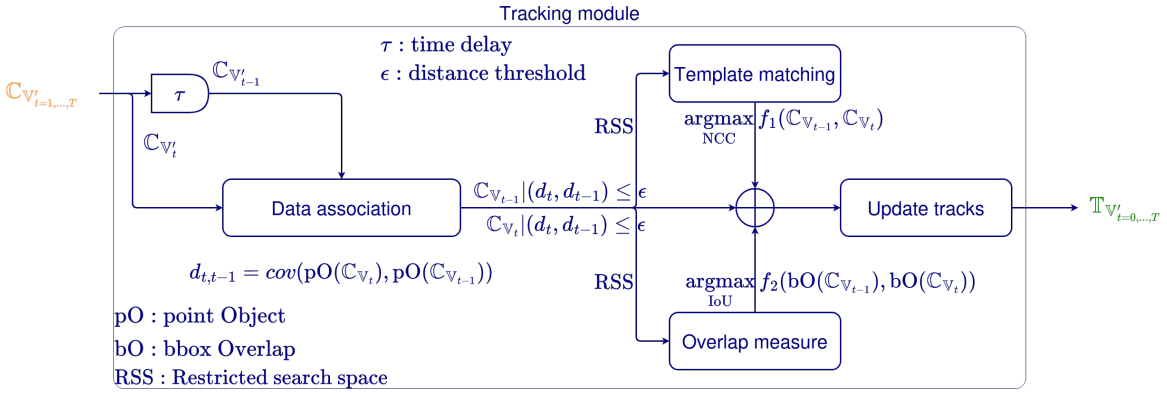

Figure 3.7: ***Structure of the meiocyte tracking module.***
Meiocyte candidates from each frame are associated in a 2-window, single hop process by restricting the search space and after that averaging over 2-measure: NCC-based template matching and IoU-based overlap ratio

- **Template matching** ensures continuity in tracking, identifying cells across sequences based on shape and positional consistency.

In tandem with EfficientDet-based ensembled detection, this advanced linking strategy constitutes a comprehensive algorithmic pipeline for cell tracking as it minimises cell switching while tracking in a crowded environment.

3.4.4.3 Biological constraints in tracking

Temporal consistency and identification

Each meiocyte's presence is validated across sequential frames, accounting for newly identified cells and those that may have become occluded or moved out of the FOV.

Movement constraints

The empirical movement constraint ensures that meiocytes are tracked within a biologically plausible search space, typically set to no more than 1.5 times the radius of their bounding box, as observed in the localized image data.

3.4.4.4 Meiocyte tracking dynamics

Temporal consistency and identification

Tracking entails maintaining a continuous record of all meiocytes identified within V'. For a given frame t, meiocytes $\mathbb{C}_{V_t}$ are compared with $\mathbb{C}_{V_{t-1}}$ to determine whether they represent

new instances or are continuations of previously tracked cells. This comparison accounts for the fact that not all meiocytes from $\mathbb{C}_{V_{t-1}}$ are guaranteed to be present in $\mathbb{C}_{V_t}$, due to potential occlusion or movement out of the field of view (FOV).

Movement constraints

To facilitate efficient and reliable tracking, the search space for meiocytes is confined based on empirical data. Within the motion-compensated videos V', meiocytes are presumed to move no more than 1.5 times the radius of their circle fitted to the bounding box from the localization stage, and movement across anthers is not considered.

3.4.4.5 Tracking methodology

Given a sequence of localized meiocytes in motion-compensated video frames $\mathbb{V}'$, the tracking module aims to follow the trajectory of meiocytes through time. For each frame t, the module identifies meiocytes $\mathbb{C}_{\mathbb{V}_t}$ and tracks these across the frames, considering the appearance of new meiocytes and the loss of existing ones due to occlusion or departure from the field of view (FOV).

Tracking meiocytes in motion-compensated video frames involves a combination of template matching, overlap analysis, and a Restricted Search Space (RSS) application:

Restricted search space through data association

To further refine the tracking process, the spatial association of meiocytes between consecutive frames $t-1$ and t is constrained within the Restricted Search Space (RSS). This is achieved by setting a distance threshold ϵ, based on the anticipated meiocyte movement. The Euclidean distance $d_{t,t-1}$ between meiocyte detections in these frames is computed as:

$$d_{t,t-1} = \|\mathbf{pO}(\mathbb{C}_{\mathbb{V}_t}) - \mathbf{pO}(\mathbb{C}_{\mathbb{V}_{t-1}})\| \tag{3.21}$$

where $\mathbf{pO}(\cdot)$ denotes the function that calculates the centroid of a meiocyte's bounding box, and $\|\cdot\|$ denotes the Euclidean distance. The RSS criterion ensures that only spatially feasible associations between meiocytes across frames are valid.

Template matching

In the template matching step, each meiocyte in the current frame $\mathbb{C}_{\mathbb{V}_t}$ is compared with meiocytes in the previous frame $\mathbb{C}_{\mathbb{V}_{t-1}}$ using Normalized Cross-Correlation (NCC). The NCC score, $\text{NCC}_{i,j}$, is calculated for meiocyte i in the current frame and meiocyte j in the previous frame. The NCC score quantifies the similarity between the two meiocytes:

$$\text{NCC}i,j = \begin{cases} f_1(\mathbb{C}^i_{\mathbb{V}_t}, \mathbb{C}^j_{\mathbb{V}_{t-1}}), & \text{if } d^{i,j}_{t,t-1} \leq \epsilon \\ 0, & \text{otherwise} \end{cases} \tag{3.22}$$

, where $d^{i,j}_{t,t-1} = |\mathbf{pO}(\mathbb{C}^i_{\mathbb{V}t}) - \mathbf{pO}(\mathbb{C}^j_{\mathbb{V}_{t-1}})|$ and $\epsilon = 1.5 \times \mathbf{r}(\mathbb{C}^i_{\mathbb{V}_t})$, $\mathbf{r}$ being the function calculating the radius of the fitted circle to the bounding box. This step is crucial for identifying potential correspondences based on appearance.

Overlap analysis

After template matching, the Intersection over Union (IoU) method assesses the spatial overlap between corresponding meiocytes. The IoU, $\text{IoU}_{i,j}$, evaluates how much the detected meiocyte regions in consecutive frames overlap:

$$\text{IoU}_{i,j} = f_2(\mathbb{C}^i_{\mathbb{V}_t}, \mathbb{C}^j_{\mathbb{V}_{t-1}}) \tag{3.23}$$

$$\text{IoU}_{i,j} = \begin{cases} f_2(\mathbb{C}^i_{\mathbb{V}_t}, \mathbb{C}^j_{\mathbb{V}_{t-1}}), & \text{if } d^{i,j}_{t,t-1} \leq \epsilon \\ 0, & \text{otherwise} \end{cases} \tag{3.24}$$

This analysis is essential for ensuring the meiocytes being compared occupy the same spatial region.

Combining template matching and overlap analysis

Finally, the tracking decision is made by combining template matching and overlap analysis results. The meiocyte j^* in the previous frame that maximizes a weighted sum of the NCC and IoU scores with meiocyte i in the current frame is selected for tracking:

$$j^* = \underset{j}{\text{argmax}}(\alpha \cdot \text{NCC}_{i,j} + \beta \cdot \text{IoU}_{i,j}) \tag{3.25}$$

where α and β are the weighting factors. This combined approach ensures robust tracking by considering both appearance and spatial congruence. Equal weighting factors are considered $\alpha = \beta = 0.5$ for combining NCC and IoU results.

Tracking update criteria

Meiocytes are tracked by updating their positions or introducing new meiocytes when no match is found. If a meiocyte from $\mathbb{C}_{\mathbb{V}_{t-1}}$ is not found in $\mathbb{C}_{\mathbb{V}_t}$, it is marked as 'invisible'. Hence, the update rule for meiocyte tracking incorporates three key criteria: cell association found, cell association not found, and the emergence of new meiocytes. This rule is mathematically articulated as follows:

- **Cell Association Found**: If a meiocyte in $\mathbb{C}_{\mathbb{V}_t}$ is associated with a meiocyte in $\mathbb{C}_{\mathbb{V}_{t-1}}$, the track is updated to include this association.
- **Cell Association Not Found**: If no association is found for a meiocyte in $\mathbb{C}_{\mathbb{V}_{t-1}}$, the meiocyte is marked as 'lost' or 'occluded'.

- **New Meiocytes Emerged**: If a meiocyte in $\mathbb{C}_{\mathbb{V}_t}$ has no association in $\mathbb{C}_{\mathbb{V}_{t-1}}$, it is added as a new track.

The mathematical representation of the tracking update rule can be defined as:

$$T^i_{\mathbb{V}_t} = \begin{cases} \mathbb{C}^{j^*}_{\mathbb{V}_{t-1}}, & \text{if } \mathrm{NCC}_{i,j^*} > 0.5 \text{ and } \mathrm{IoU}_{i,j^*} > 0.45, \\ \mathbb{C}^{i}_{\mathbb{V}_{t-1}}, & \text{if } i \text{ is unassociated and } \mathrm{NCC}_{i,j^*} \leq 0.5 \text{ and } \mathrm{IoU}_{i,j^*} \leq 0.45, \\ T^{i}_{\mathbb{V}_{t-1}}, & \text{if } i \text{ becomes invisible (no association found).} \end{cases} \tag{3.26}$$

3.4.4.6 Tracking evaluation

The tracking evaluation is done by two widely used metrics in multiple object tracking [62]: MOTA (Multiple Object Tracking Accuracy) and MOTP (Multiple Object Tracking Precision). These metrics are crucial in evaluating the performance of multiple object-tracking algorithms in video processing and computer vision.

MOTA(Multiple Object Tracking Accuracy) evaluates the overall tracking accuracy, accounting for three types of errors:

- False Positives (FP): Detections not associated with ground-truth objects.
- False Negatives (FN): Ground-truth objects that were undetected.
- Identity Switches (IDSW): Instances where a tracked trajectory changes its matched ground-truth identity.

The formula for MOTA is given by:

$$\mathrm{MOTA} = 1 - \frac{\mathrm{FP} + \mathrm{FN} + \mathrm{IDSW}}{\text{Total Ground Truth Objects}} \tag{3.27}$$

A higher MOTA score indicates better tracking performance. The score can range up to 1 (or 100%), and can also be negative if the number of errors exceeds the number of ground truth objects.

MOTP (Multiple Object Tracking Precision) measures the precision of object localization. It calculates the average distance between objects' predicted and ground-truth positions. The formula for MOTP is:

$$\mathrm{MOTP} = \frac{\sum_{t,i} d_{t,i}}{\sum_t c_t} \tag{3.28}$$

, where:

- $d_{t,i}$ is the distance between the object's predicted and ground truth position i at time t.
- c_t is the number of correctly matched pairs at time t.

A lower MOTP value indicates higher precision in tracking, as the predicted positions are closer to the ground truth.

MOTA and MOTP are essential for evaluating different performance aspects of a tracking system. MOTA focuses on the ability to keep consistent tracks of objects, while MOTP assesses the spatial precision of the tracking. The use of MOTA and MOTP as our evaluation metrics is specifically chosen for their ability to accurately reflect the performance of our tracking system in a biological context. These metrics help us assess not just the accuracy of tracking, but also the precision in capturing the nuanced movements and changes of meiocytes.

3.4.4.7 Tracking output

The output of the tracking module is a set $T_{\mathbb{V}_t}$ for each frame t, detailing the tracked meiocytes, their updated positions, new identifications, and visibility statuses. This creates a comprehensive temporal profile of meiocyte dynamics within the video sequence. The tracking module generates an updated collection, $T_{\mathbb{V}'_t}$, for each frame t. This collection captures the tracked meiocytes, detailing their trajectories and current visibility within the video. Such meticulous tracking is crucial for the downstream components of the aMP pipeline, as it ensures precise monitoring of meiocyte activity over time. The data provided by this module is essential for thoroughly examining meiocyte progression and their developmental stages across the video sequence.

3.4.5 The meiocyte classification module

The classification module within the aMP pipeline is designed to assign a state indicative vector, denoted as siv_t, to each tracked meiocyte in $\mathbb{T}_{\mathbb{V}_t}$. As elaborated in the dataset description, this involves categorizing the cropped and tracked meiocyte image regions according to a set of five interrelated labels: Cell Shape (CS), Nucleus position (NP), Nucleolus position (NoP), Microtubule Array (MT), and Chromosome (RC) dynamics. The operational flow of this classification module is illustrated in Fig. 3.8.

3.4.5.1 Multi-class and multi-label classification: a literature review

Multi-class and multi-label classification has emerged as a significant area in machine learning, addressing complex problems where each instance may be associated with multiple labels. Unlike traditional classification tasks where each instance is linked to a single label, multi-label classification (MLC) deals with instances that can simultaneously belong to many classes [63].

The importance of MLC is underscored by its wide applicability in diverse fields such as text categorization, image and video classification, and bioinformatics [64]. Bioinformatics, for example, is used for predicting gene functions where each gene may be associated with multiple functions. In multimedia processing, an image might be tagged with multiple labels like 'beach', 'sunset', and 'vacation' [65].

Madjarov et al. (2012) made a crucial contribution to the field by conducting an extensive experimental comparison of MLC methods. They evaluated 26 methods across 42 benchmark datasets using 20 evaluation measures, providing a comprehensive understanding of the strengths and weaknesses of various approaches [66].

MLC methodologies can be broadly categorized into problem transformation and algorithm adaptation [67]. Problem transformation methods like Label Powerset (LP) and Pruned Sets (PSt) transform the multi-label problem into single-label datasets, which are then addressed using conventional classifiers. Algorithm adaptation methods, on the other hand, modify existing algorithms to handle multiple labels directly, as seen in approaches like Classifier Chains and AdaBoost for MLC [68, 69].

Classifier Chains (CC), a novel approach, demonstrates the ability to exploit correlations among labels, a critical aspect often overlooked in traditional methods [68]. In CC, classifiers are linked in a chain, with each classifier dealing with the binary relevance problem for a single label and including the predictions of all previous classifiers in the chain as features.

The process flow in MLC typically involves the following stages: data preprocessing, feature selection, classifier training, and evaluation. Effective feature selection is crucial for handling the high dimensionality in MLC problems. The evaluation of MLC methods requires specialized metrics, such as Hamming loss, precision, recall, and F1 score, as these methods must account for the accuracy of predicting multiple labels simultaneously [63].

Studying multi-class and multi-label classification is vital for advancing machine learning applications in complex real-world scenarios. The continuous development of new methods and comparative studies enhances our understanding and capability to handle multi-label datasets effectively. As this field evolves, it is imperative to consider both the predictive performance and computational efficiency of these methods, particularly in large-scale data.

3.4.5.2 Rationale for employing average ensembles in multi-class and multi-label classification

Building upon the previously established EfficientNet backbone, our approach to multi-class and multi-label classification leverages the concept of average ensembles of different classifier networks. Average ensembles combine the predictive power of multiple models, reducing variance and improving generalization. This approach aligns with the findings of Madjarov et al. (2012), who underscore the significance of ensemble methods in enhancing classification performance across diverse datasets [66].

The choice to use average ensembles is further strengthened by our use of the EfficientNet backbone. EfficientNet, known for its balance between accuracy and computational efficiency, provides a robust feature extraction base for our classifiers. By averaging the outputs of various classifiers built upon this backbone, we can exploit the complementary strengths of different learning algorithms, a strategy echoed in the works of Read et al. (2011) regarding Classifier Chains and AdaBoost [68, 69].

Multi-label classification inherently involves complex label dependencies and varying degrees of label imbalance. By integrating predictions from multiple models, average ensembles offer a nuanced approach to capturing these dependencies. This method resonates with the algorithm adaptation strategies discussed by Zhou and Zhang (2006), whose goal is to directly accommodate the multi-label nature of the problem [67].

In line with the principles outlined by Zhang and Zhou (2014), our ensemble approach aims to enhance predictive performance while addressing the challenges of multi-label classification,

such as high dimensionality and the need for specialized evaluation metrics [63]. By averaging classifiers' outputs, we aim to balance precision and recall, which is crucial in multi-label settings.

Average ensembling on softmax probabilities

Let's assume we have N models, and each model i provides a softmax output vector $\mathbf{p}^{(i)} = \left[p_1^{(i)}, p_2^{(i)}, \ldots, p_C^{(i)}\right]$ for a given input, where C is the number of classes. The softmax output for each class j in model i is denoted by $p_j^{(i)}$, which is the predicted probability of the input belonging to class j.

The average ensemble softmax output $\mathbf{p}^{\text{ens}}$ is calculated as the average of the softmax outputs of all models for each class:

$$\mathbf{p}^{\text{ens}} = [\bar{p}_1, \bar{p}_2, \ldots, \bar{p}_C] \tag{3.29}$$

$$\text{where } \bar{p}_j = \frac{1}{N} \sum_{i=1}^{N} p_j^{(i)} \tag{3.30}$$

In this approach, the average ensemble softmax output, $\mathbf{p}^{\text{ens}}$, is a vector where each element $\bar{p}_j$ represents the averaged probability of the input belonging to class j across all models. This method effectively pools the predictions from each model, where each model has an equal contribution to the final decision. It helps reduce the prediction variance, as individual model biases or errors are likely to be averaged out, potentially leading to more robust and generalizable results.

By averaging across models, we can leverage the strengths of different models while mitigating the weaknesses of individual ones. This is particularly useful in scenarios where different models capture different aspects of the data or have varying degrees of sensitivity to certain features or classes. The average ensemble method is straightforward yet powerful, especially when the constituent models are diverse and complementary.

3.4.5.3 Training methodology

In this study, we trained an ensemble of EfficientNet B0, B1, and B2 models, leveraging their complementary strengths to address the multi-class and multi-label classification task. The choice of EfficientNet models, known for their scalability and efficiency, aligns with our objective to balance computational resources and model performance.

The training of these models employed the Adam optimizer [52]. For the loss function, we utilized categorical cross-entropy, modified to incorporate inter-class and inter-label weightage. Given the dataset's inherent class imbalance, where certain meiotic states appear more frequently, the loss function considers both the label and class frequency information, allowing the model to pay more attention to less frequent but significant classes, enhancing the model's sensitivity to minority classes. The mathematical representation of the loss function is given

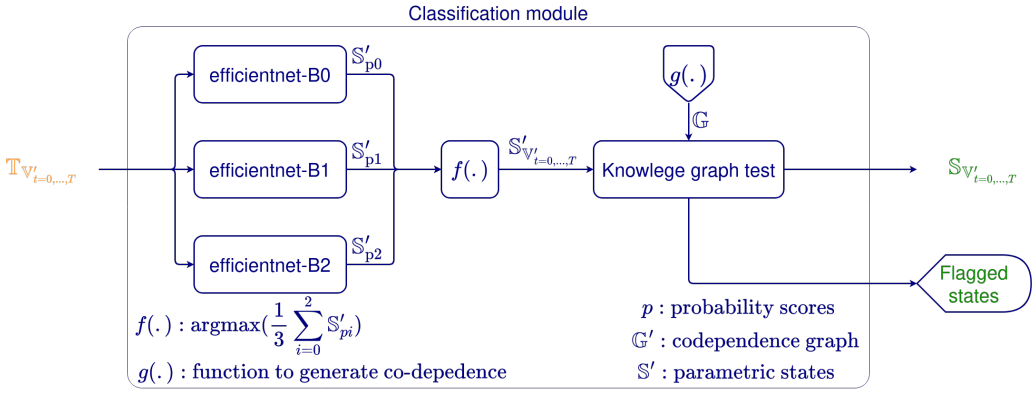

Figure 3.8: ***Structure of the classification module.***
The tracked meiocytes are classified in an ensembled-averaged method, which is post-processed and automatically corrected and flagged by a knowledge graph test to provide the parametric states of meiocyte.

as:

$$\text{Total loss} = \sum_{\text{sip}}^{[\text{CS,NP,NoP,MT,RC}]} w_{\text{sip}} \cdot \text{sip}_{\text{WCC}} \tag{3.31}$$

$$\text{sip}_{\text{WCC}} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} w_j \cdot y_{ij} \cdot \log(p_{ij}), \tag{3.32}$$

where N is the total number of samples, C is the number of classes, y_{ij} is the binary indicator (0 or 1) if class label j is the correct classification for sample i, p_{ij} is the predicted probability of sample i belonging to class j, and w_j is the weight assigned to class j. The term sip_{WCC} represents the weighted categorical cross-entropy for a specific label sip, and w_{sip} is the weight reflecting the relative importance of each label sip in the overall loss function.

We used categorical accuracy as the metric for model performance evaluation. This metric is a standard choice in multi-class classification scenarios and directly measures how often the model predicts the correct label [70]. We employed an exponential learning rate decay to refine our training process further. This strategy helps converge to the optimal solution more efficiently by gradually reducing the learning rate, thus preventing oscillations near the global minimum.

3.4.5.4 Inference pipeline

Our meiocyte classification system employs an inference pipeline that integrates a robust classification module with a knowledge graph test to categorize cells into five critical parameters: cell shape, nucleus position, nucleolus position, chromatin, and microtubule array. The pipeline functions as follows:

Upon receiving input tensor $\mathbb{T}_{v_{0,\ldots,T}}$, which comprises the visual data for a given meiocyte over time T, the system directs the input through an ensemble of EfficientNet models—specifically, EfficientNet-B0, EfficientNet-B1, and EfficientNet-B2. Each model is tasked with extracting features and predicting the state vectors S'_{p0}, S'_{p1}, and S'_{p2}, respectively, for each parameter. These vectors represent the initial predictions for the five parameters of interest.

The classification module then applies the function $f(\cdot)$, defined as:

$$f(\cdot) = \text{argmax}\left(\frac{1}{3}\sum_{i=0}^{2} S'_{pi}\right)$$

This function calculates the argmax over the averaged predictions from the ensemble, yielding a consensus prediction $S'_{v_{0,\ldots,T}}$ for each parameter across all time points.

Subsequently, the knowledge graph test is initiated using the function $g(\cdot)$, which incorporates domain-specific knowledge through a codified graph G to test the initial predictions. This process ensures that the final predictions $S_{v_{0,\ldots,T}}$ are not only based on the ensemble's outputs but are also refined through biological constraints and relationships defined within the knowledge graph.

The output of the knowledge graph test is then used to derive the final predictions $S_{v_{0,\ldots,T}}$, which are classified into their respective classes for each of the five parameters. Any states flagged as inconsistent or highly improbable given the biological context are highlighted as "Flagged states," prompting further investigation or manual review.

Integrating deep learning with domain-specific knowledge, this inference pipeline provides a priory-based approach to meiocyte classification. It demonstrates the power of machine learning when combined with expert knowledge, leading to a more accurate and contextually relevant classification system for complex biological data.

Introducing the knowledge graph

The primary biological query of our study centres on the classification of meiotic stages and associated cellular parameters. Given the dynamic nature of meiosis, it is crucial that the classification block not only identifies the stages but also corroborates them against the expected biological progression. To this end, a post-processing step has been implemented that employs a knowledge graph, essentially an adjacency matrix, reflecting the theoretical underpinnings of meiotic processes.

This knowledge graph is constructed to encapsulate two fundamental phenomena of meiosis: the monotonic progression of meiotic stages and the intricate interdependencies among various cellular parameters, such as cell shape, chromatin structure, nucleus position, nucleolus position and microtubular array. The monotonic nature of the stages of meiosis ensures that the progression follows a predefined sequence that does not revert. Meanwhile, the interdependencies among cellular parameters are informed by the tight regulation of meiotic processes, where changes in one aspect often correlate with transformations in others.

Within our post-processing framework, the knowledge graph serves a dual purpose. First, it acts as a biological filter, ensuring that the classification outcomes are consistent with the

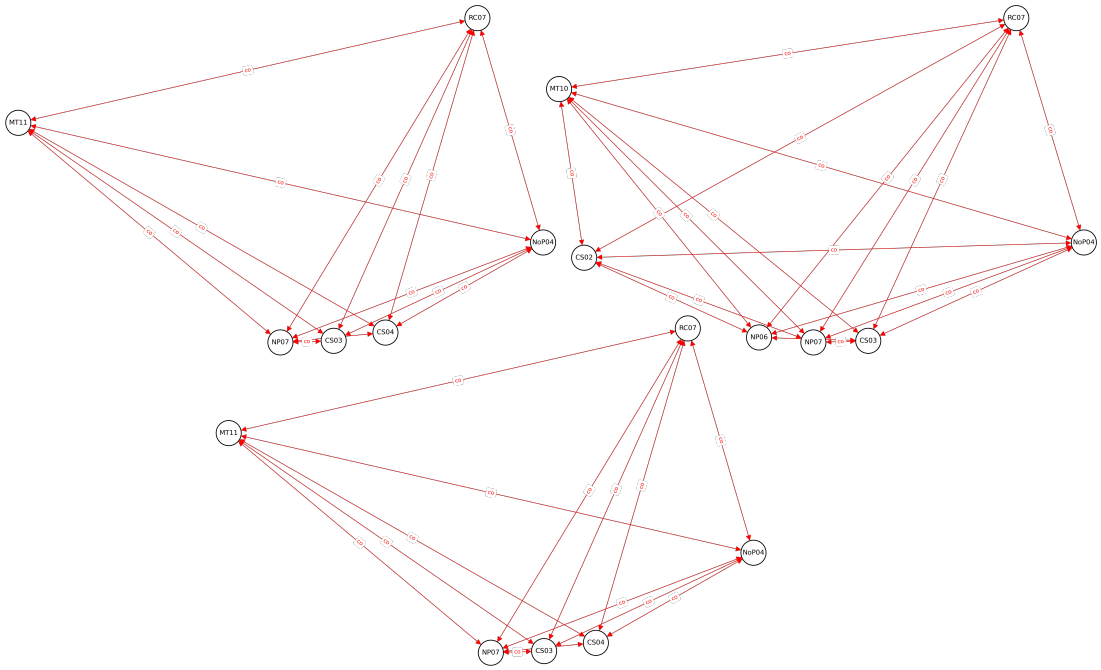

Figure 3.9: ***Excerpt from Knowledge graph:***
The meiosis knowledge graph represents the interdependence nature of the 5 cellular parameters shown in Fig 2.4. The knowledge graph is 2-layered: 1. the progressive directional intra-parametric path representing the monotonic state increment, along with the option of staying in the same state, 2. the bidirectional inter-parametric dependence of the cell parameters.

linear progression of meiotic stages. Any deviations from this progression, potentially due to experimental artefacts such as shifts in the z-stack plane attributed to the flower's growth or the translocation of nuclei during the second division of meiosis—are rigorously scrutinized. Second, the graph assists in auto-correction classifications, leveraging the interconnections between cellular parameters to adjust or flag results that fall outside expected patterns.

Classifications that appear incongruent with the established progression, as interpreted by the knowledge graph, are subjected to adjustment algorithms or marked as ambiguous. This marking is contingent on a defined threshold that quantifies the classification confidence relative to the biological expectations set by the graph. By integrating this threshold-based system, we impose a rigorous standard that upholds biological validity while accommodating the inherent variability of biological systems.

Given the classifier's output for a meiocyte image as the state vector

$$S'_{i,k} = [S'_{\mathrm{CS}}, S'_{\mathrm{NP}}, S'_{\mathrm{NoP}}, S'_{\mathrm{MT}}, S'_{\mathrm{RC}}]_i,$$

where each S'_i represents the predicted class for the respective parameter (cell shape, nucleus position, nucleolus position, microtubule array, and reticulate chromatin), and the knowledge graph's adjacency matrix A, which captures the expected transitions and interdependencies between different meiotic stages and cellular parameters, the post-processing steps can be

described as follows:

1. **Validation against knowledge graph:** Each transition in the predicted sequence is validated against A to ensure biological plausibility:

$$L1(S'_i, S'_{i+1}) = \sum_k |S'_{ik} - S'_{i+1,k}| \tag{3.33}$$

   where k iterates over all the parameters within the state vector.

   The threshold $\delta = 1$ is a predefined value that represents the maximum allowable difference between two valid consecutive states which conserves the monotonic meiosis progression. The validation against the adjacency matrix A and the $L1$ norm is incorporated into the validation function:

$$\text{Valid}(S'_i, S'_{i+1}) = \begin{cases} \text{true} & \text{if } A_{S'_i, S'_{i+1}} = 1 \text{ and } -\delta \leq L1(S'_i, S'_{i+1}) \leq \delta, \\ \text{false} & \text{otherwise.} \end{cases} \tag{3.34}$$

2. **Auto-Correction:** For any sequence where $\text{Valid}(S'_i, S'_{i+1})$ is false, an auto-correction is attempted by finding a more probable state given the previous classifications and the knowledge graph:

$$S'_{i+1} = \underset{j}{\text{argmax}}\, A_{S'_i, j}. \tag{3.35}$$

   This step replaces the less probable state with a more probable one according to A.

3. **Ambiguity Handling:** If $A_{S'_i, S'_{i+1}}$ is below a certain confidence level but no better alternative is found, the state is marked as ambiguous:

$$S''_i = \begin{cases} S'_i & \text{if } A_{S'_i, S'_{i+1}} \geq \delta, \\ \text{ambiguous} & \text{otherwise.} \end{cases} \tag{3.36}$$

   Here, $S''_i \in \mathbb{S}''$ represents the adjusted state vector after post-processing.

4. **Final State Vector:** The final state vector S'' is then compiled, which contains original predictions, corrected predictions, or ambiguities that require further investigation.

In conclusion, the classification module is a central element of the aMP pipeline, exploiting the capabilities of EfficientNet models through an ensemble approach to accurately categorize meiocyte states. The integration of a knowledge graph with the classification process ensures that predictions are not solely based on image data but are also informed by biological context, resulting in a refined list of state indicative vectors $\mathbb{S}''_t$. These vectors correspond to the list of meiocytes $\mathbb{C}_t$, offering a classification that synergizes immediate image-based evidence with the broader biological narrative of meiotic progression.

The forthcoming evaluation section is tasked with thoroughly analysing this module's performance. It will provide a critical assessment of the classifications' precision and recall, affirming the robustness of our approach or highlighting opportunities for enhancement. Thus, the evaluation will substantiate the classification module's contribution to our understanding of meiotic events, as reflected in the state indicative vectors $\mathbb{S}''_t$, which encapsulate the complex interplay between observed cellular features and the expected biological sequence of meiosis.

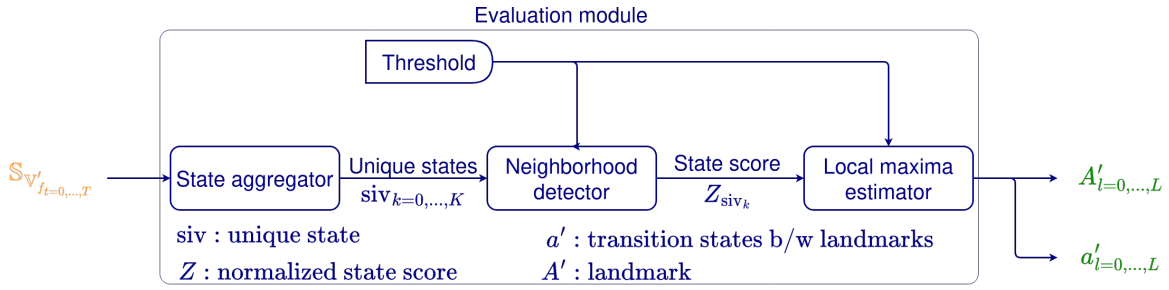

Figure 3.10: ***Structure of the evaluation module.***
The vectored meiocytes undergo a state aggregation to determine their unique states and occurrence count. A biologically inspired threshold is applied for the state score to find the neighbours of uniquely determined states and thereby find local maxima (landmark) are identified.

3.5 The evaluation module

Following the detailed classification of meiocytes, the evaluation module plays a pivotal role in quantitatively analyzing meiotic progression. Utilizing the state-indicative vectors obtained from the classification module, this component systematically identifies key landmark states and computes their transitions. This process effectively elucidates the nuanced timeline of meiosis in Arabidopsis thaliana, integral to understanding the cellular dynamics of this crucial biological process.

Evaluating meiotic progression is a pivotal component in understanding the cellular division intricacies. The aMP pipeline introduces an evaluation module, as depicted in Figure 3.10, designed to identify meiotic landmark states within the dataset. Landmark states are defined by their prevalence in the dataset, based on the state indicative vectors (siv) representation of the meiocytes.

The meiotic landmark system is employed to quantify and analyze the stages of meiosis by identifying distinct structural and morphological 'landmarks' that occur during this process. This system is particularly useful in live cell imaging studies, which aid in the detailed observation and timing of meiotic events. The landmark system involves identifying key stages or 'landmarks' in meiosis. These landmarks are specific, observable changes in the structure and organization of cells and chromosomes at different stages of meiosis. By tracking these landmarks, scientists can achieve a high temporal resolution in their study of meiosis. This means they can determine how long each stage of meiosis lasts and identify any deviations from the normal process, particularly in mutant or genetically altered specimens. The landmark system is particularly useful in analyzing meiotic mutants. By comparing the occurrence and timing of these landmarks in mutant cells to those in normal cells, we can identify how certain genetic changes affect the progression of meiosis. Understanding when and how these landmarks occur provides scientists with insights into regulating meiosis, including the mechanisms that control chromosome pairing, recombination, and segregation-critical processes for ensuring genetic diversity and stability.

3.5.1 Mathematical framework for landmark estimation

The evaluation of meiotic progression in the aMP pipeline starts with annotating unique state indicative vectors (SIVs) from the classification output, delineating a distinct set of states within the dataset. Each unique SIV undergoes a computational analysis to establish its relationship with other states.

$$siv = \text{Unique}\{siv_1, siv_2, \ldots, siv_n\} \tag{3.37}$$

The process begins with the calculation of distances between unique SIV pairs. This is achieved by summing the absolute differences, analogous to Eq. 3.33 in each of their constituent parameters, resulting in a matrix that captures the dissimilarities between every possible pair of states.

$$\text{params} = \text{CS, NP, NoP, MT, RC} \tag{3.38}$$

$$L(siv_k, siv_j) = \sum_{\text{params}} |siv_{k_{\text{params}}} - siv_{j_{\text{params}}}| \tag{3.39}$$

Next, neighbouring states for each SIV are identified. A neighbouring state varies minimally (by a distance of no more than one unit) from the reference SIV in only one component. This adjacency criterion is crucial for understanding the transition potential between different meiotic states.

$$ns_k = \{siv_j | -\delta \leq L(siv_k, siv_j) \leq \delta\}, \tag{3.40}$$

where the lone constraint $\delta = 1$ is set, similar to Eq. 3.33.

To evaluate the prominence of each unique state in the dataset, the mean and standard deviation of occurrences for all its neighbouring states are calculated. This statistical analysis provides insight into each state's relative frequency and variance within the context of its immediate neighbours.

$$\mu_{siv_k} = \frac{1}{N_k} \sum_{j=1}^{N_k} c_{ns_{kj}} \tag{3.41}$$

$$\sigma_{siv_k} = \sqrt{\frac{1}{N_k - 1} \sum_{j=1}^{N_k} (c_{ns_{kj}} - \mu_{siv_k})^2} \tag{3.42}$$

$$Z_{siv_k} = \frac{c_{siv_k} - \mu_{siv_k}}{\sigma_{siv_k}}, \tag{3.43}$$

the critical step in the landmark determination process is the computation of Z-scores for each unique state. These scores are derived by comparing the frequency of a unique state c_{siv_k} with the average frequency of its neighbouring states μ_{siv_k}, normalized by the standard deviation σ_{siv_k}. A unique state with a Z-score $\geq \delta$ signifies a landmark state A_l, indicating a higher prevalence compared to its neighbours and marking it as a significant point in the meiotic progression, which can be formulated as:

$$A_l = siv_k|_{Z_{siv_k} \geq \delta} \tag{3.44}$$

The final part of the evaluation involves the construction of the meiotic timeline. This is

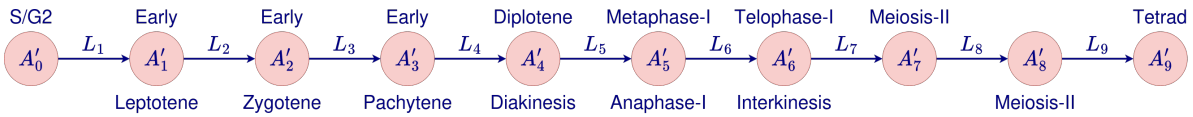

Figure 3.11: Landmark points and the transition pathway

achieved by mapping the transitions between these landmark states and calculating the time intervals for progression from one landmark to another. This temporal analysis provides a detailed overview of the meiotic stages, capturing the dynamic nature of cellular development during meiosis. A 'transition' represents the movement from one landmark state to the next, signifying a definitive progression in meiosis. The 'time of progression' refers to the duration needed for a meiocyte to transition to the beginning of the next landmark state. For each pair of landmarks $L_l = (A_{l-1}, A_l]$ (Fig. 3.11), the transition-time T_l is computed for each live imaging video as follows:

$$T_l = t_{A_l|S_V} \tag{3.45}$$

$$t_{A_l|S_V} = \left[c_{a_l \cup A_l} \times T_{s|V}\right]_{A_{l-2} \in \mathbb{S}_\mathbb{V}}, \tag{3.46}$$

where T_s is the sampling time for video $\mathbb{V}$, a_l represents the transition states within (A_{l-1}, A_l), and A_l are the identified landmarks. The meiotic timeline is constructed as a piece-wise representation between landmarks to encompass the genetic diversity and overlapping nature present in the videos. This methodology considers the variability in the start and end times of the imaging sessions.

3.6 (Hyper)parameter choices in aMP-kit modules

In the aMP pipeline, various modules necessitate distinct neural network models and hyperparameter configurations. For instance, the segmentation module integrates U2-Net and EfficientNet architectures, utilizing Focal Tversky loss and an Adam optimizer with a learning rate initially set at $1e-4$. The module also employs early stopping, adjusting batch sizes based on computational capacity.

The stabilization module applies a translation-only motion model, with a restricted search range tailored to the expected anther movement. It employs a custom centroid tracking-based optimization algorithm. In contrast, the localization module leverages single-shot detectors using EfficientDet backbones (D0, D1, D2), with Smooth L1 loss for bounding box accuracy and focal loss for objectness. Its optimizer, Adam, is set with a learning rate of $1e-3$, and non-maximum suppression is applied with an IoU threshold of 0.5.

Tracking is facilitated through Normalized Cross-Correlation and Intersection over Union, with a movement constraint set at 1.5 times the meiocyte's bounding box radius. The classification module utilises the EfficientNet backbones (B0, B1, B2) with fully connected heads. The loss function is a weighted cross-entropy to account for class imbalances, and the models are trained using Adam with a learning rate beginning at 1e-4. An averaging ensemble strategy is applied, and performance is evaluated based on class-balanced accuracy and F1 score.

Time progression analysis involves identifying landmark states using frequency and Z-score

calculations of state-indicative vectors. This process measures the intervals between landmark states, which is crucial for understanding the timeline of meiotic progression.

These hyperparameters are tuned by grid search, each optimized for the specific task within the pipeline, be it segmentation, stabilization, localization, tracking, classification, or time progression analysis. The iterative refinement of these parameters, tabulate in Table. 3.1, validated through testing, underscores the pipeline's capacity to analyze meiosis progression in *Arabidopsis thaliana* effectively.

Module	Parameter	Value/Range
Segmentation	Loss function	Focal Tversky loss
	Optimizer	Adam with decay
	Learning rate	1e-4 with decay
	Batch size	8
	Epochs	Early stopping
Stabilization	Motion model	Rotation and Translation
Localization	Loss function	Smooth L1, Focal loss
	Optimizer	Adam with decay
	Learning rate	1e-3 with decay
	Batch size	8
	NMS threshold	IoU@0.5
Tracking	Distance threshold	1.5× bbox radius
Classification	Loss function	Weighted cross-entropy
	Optimizer	Adam with decay
	Learning rate	1e-4 with adjustment
	Batch size	8
	Ensemble strategy	Averaging ensemble
	Validation metrics	Accuracy, F1 score
Time progression	Landmark algorithm	Frequency, Z-score

Table 3.1: *(Hyper)parameter choices in aMP-kit modules*

3.7 Implementation Details

Implementing the aMP pipeline leveraged the TensorFlow framework for its robust support in deep learning capabilities. A dual hardware setup provided the computational backbone of this implementation. The central processing unit (CPU), an Intel Core i5 8600K with a base frequency of 3.6GHz and 32GB RAM, ensured efficient data handling and preprocessing. Complementing this, the graphics processing unit (GPU), an NVIDIA RTX 2080Ti equipped with 11GB of memory, offered the computational power necessary to train deep learning models. This combination of CPU and GPU provided an optimal balance between processing speed and memory capacity, essential for handling the computationally intensive tasks involved in image-based analysis within the aMP pipeline.

This chapter has detailed the methodology employed in the Automated Meiosis Progression (aMP) pipeline, explaining the intricate processes and tools used in analyzing meiosis progression in Arabidopsis thaliana. The next chapter will present the results obtained using these

methods. It will offer insights into the efficacy of the pipeline in extracting and analyzing crucial data and evaluate how these findings contribute to our understanding of meiotic progression. This chapter aims to bridge the theoretical approaches outlined in Chapter 3 with practical, data-driven results, showcasing the real-world application and impact of the aMP pipeline.

Chapter 4

Framework evaluation

In this chapter, we critically evaluate the components of our framework to demonstrate its effectiveness in analyzing meiocytes through the stages of meiosis, ultimately facilitating the creation of a comprehensive timeline analysis. Our evaluation methodically addresses each key phase of the framework, starting with the segmentation of anthers to ensure stability, followed by the precise localization of meiocytes across video frames, and concluding with the tracking and staging of these cells throughout their developmental journey.

The timeline evaluation is systematically divided into three distinct sections. Initially, we focus on the significance of our approach in accurately grouping and staging meiotic events, providing a foundation for subsequent analyses. This is complemented by a detailed comparison with existing studies in the field, highlighting the association our framework offers with previously published results. Finally, applying our methodology to tetraploids underlines the framework's utility in revealing the meiosis timeline of tetraploids compared to diploids.

4.1 Results of anther segmentation and stabilization

The U^2-Net deep learning architecture was employed for segmenting anthers from high-resolution plant images. Our results demonstrate significant success in this task, as evidenced by the achieved dice score (an accuracy measure in segmentation tasks) in the test set of 88.9%. This accuracy level indicates the model's robustness in distinguishing anthers from the complex background and the tapetum structures. Precision, a measure of the model's ability to correctly identify positive anther instances, was recorded at an impressive 87.9% for the test set. The recall rate, indicating the model's capacity to identify all relevant anther instances, stood at 91.1% in the test set. The F1 score, which balances precision and recall, was calculated to be 88.9% for the test set, reflecting the model's overall efficiency in anther segmentation. Furthermore, the Intersection over Union (IoU) for the test set was notably high at 80.8%, showcasing the model's precise delineation of anther boundaries.

Continuing from the results presented earlier, the similarity between the Dice Coefficient and the F1 Score is worth discussing. Both metrics focus on the balance between precision and recall, which is crucial for evaluating accuracy in binary segmentation tasks. The Dice Coefficient and the F1 Score essentially convey the same information. In our results, the Dice Coefficient

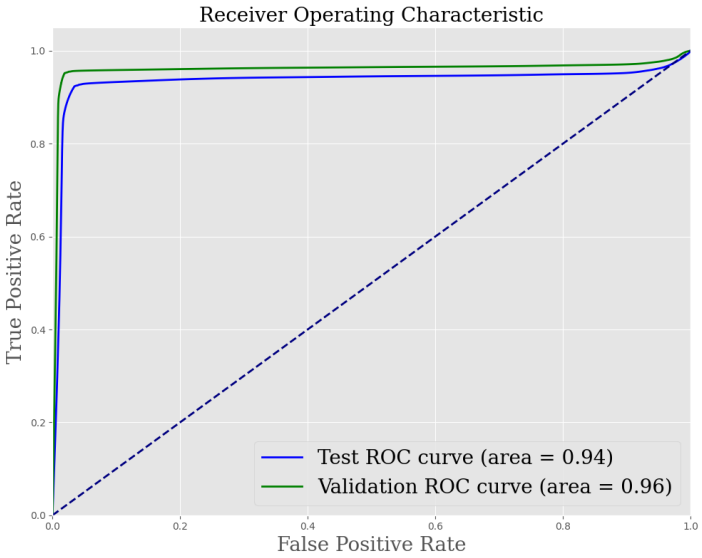

Figure 4.1: Receiver operating characteristics of the anther segmentation

for the test set stood at 88.9%, and the F1 Score was also 88.9%, reinforcing their conceptual similarity and practical interchangeability in binary classification tasks like image segmentation. This similarity underlines the consistency and reliability of the metrics used in evaluating our U^2-Net architecture's performance in segmenting anthers and is tabulated in Table 4.1.

The ROC curves in Fig. 4.1 further substantiate the efficacy of the U^2-Net architecture in segmenting anthers from high-resolution plant images. The test set yielded an AUC of 0.94, indicating a high true positive rate with a relatively low false positive rate across various thresholds, a testament to the model's accuracy. Remarkably, the validation set presented an even higher AUC of 0.96, demonstrating exceptional model reliability during the tuning phase. These AUC values indicate the model's robust discriminating power in distinguishing anther pixels from background textures. Such high AUC values and the previously discussed Dice Coefficient and F1 Score confirm the model's precision in segmenting anthers with minimal error, reinforcing confidence in the model's performance metrics.

Visual representations of these results further corroborate the quantitative findings. Sample segmentation outputs are presented in Fig. 4.2, which showcases images selected from diverse

Metric	Validation Set	Test Set
IoU	0.8897	0.8075
Dice Coefficient	0.9399	0.8890
Precision	0.9371	0.8798
Recall	0.9486	0.9108
F1 Score	0.9399	0.8890

Table 4.1: *Evaluation Metrics for Validation and Test Sets.*
Quantitative evaluation was based on the Dice coefficient to quantify the overlap of the ground truth segmentation masks and the segmentation results of the segmentation network (value range of the Dice coefficient: [0,1]; 1: perfect overlap). For the validation videos of the Arabidopsis thaliana dataset 1, i.e. the videos used for hyperparameter optimization, the mean Dice coefficient was 93.9%. For the testing videos of dataset 1, it was 88.9%

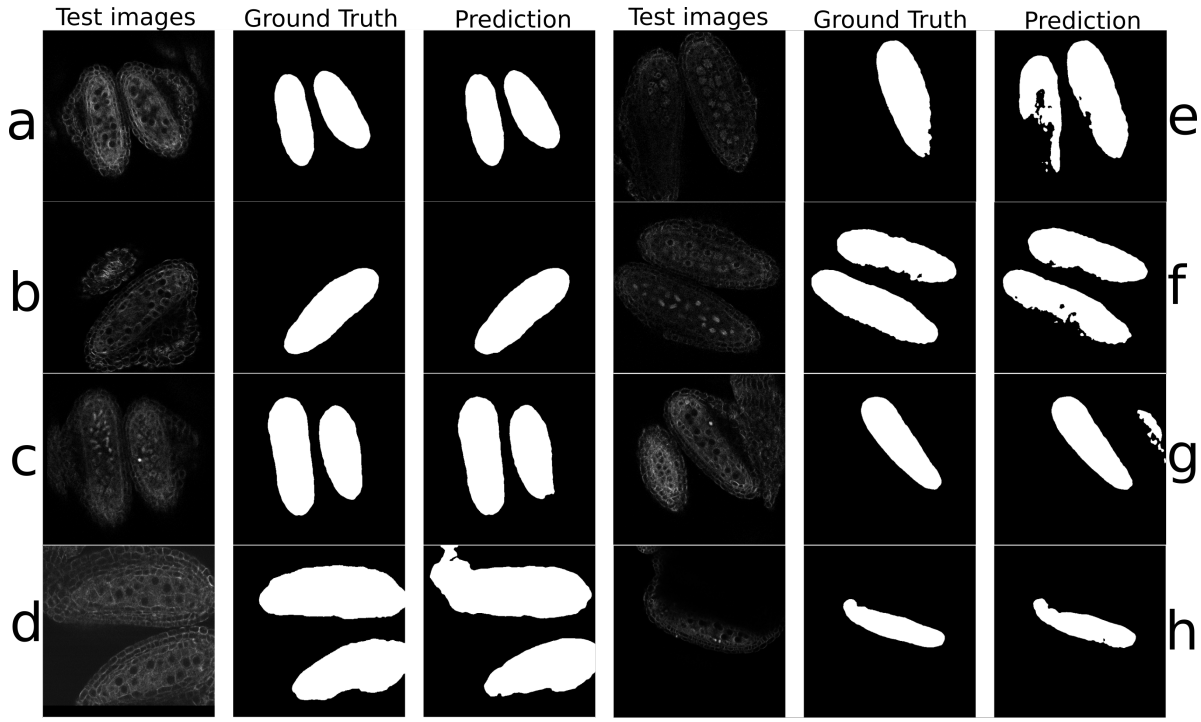

Figure 4.2: Comparative anther segmentation analysis on live imaging dataset. The left column presents the original images of anthers; the centre column shows the corresponding ground truth masks and the right column displays the predicted segmentation masks. The images are selected to demonstrate the model's performance across various challenging conditions, highlighting its capabilities and occasional limitations in segmenting anthers with varying degrees of contrast, orientation and structural overlap with other floral elements.

scenarios within our test dataset. These representations illustrate the model's adeptness in accurately delineating anthers across various conditions, such as different lighting, orientations, and instances where anthers overlap with other floral structures. Notably, the segmentation maps demonstrate a minimal occurrence of false positives, emphasizing the model's precision in boundary delineation - a testament to the high Dice Coefficients and F1 Scores previously discussed.

Despite the overall high performance, our analysis also identified certain limitations. Specifically, there were occasional challenges in segmenting anthers in cases of significant overlap with surrounding flora, leading to minor under-segmentation artefacts. Similarly, in some images where the contrast between the anthers and the background was exceptionally low, the model exhibited tendencies toward slight over-segmentation. These instances, however, were relatively infrequent and did not substantially detract from the model's overall efficacy, as reflected in the high IoU scores and the robust AUC values indicated by the ROC curves for both the validation and test sets. When combined with the quantitative metrics, the visual insights affirm the model's effectiveness and highlight areas for future optimization to enhance its segmentation performance further.

In a comparative analysis with existing segmentation methodologies with FIJI/ImageJ, employing different case-specific thresholding and edge-detection algorithms, the U^2-Net model demonstrated highly superior time efficiency and accuracy. This was particularly evident in its handling of complex and varied backgrounds, offering a more nuanced and precise segmenta-

tion. Table. 4.2 briefly summarises the speed of our segmentation operation compared to Fiji, where the speed of operation varies with user based on the pre-processing steps required to perform and setting the threshold parameters specific to the acquired images.

Metric	FIJI	Our method
speed	User dependent	10.15 fps

Table 4.2: Quantitative comparison

4.1.1 Anther stabilization through affine transformation

Further to the segmentation results, we addressed the crucial aspect of anther stabilization, which relied on the segmented anther masks generated by the U^2-Net model. This stabilization technique aimed to correct positional variances of anthers across sequential image frames, ensuring consistent orientation and positioning for in-depth analysis. The data from the Train, Validation, and Test datasets clearly illustrate the significant impact of the stabilization process on both average displacement and standard deviation. A notable reduction in the mean values of 'Stabilized Avg Displacement' and 'Stabilized Std Deviation' compared to the original values across all datasets indicates the effectiveness of the stabilization technique. This reduction is crucial as it suggests a successful minimization of variations and movements within the data, a primary objective of stabilization processes.

Dataset-Specific Observations

In the Train Dataset, the range of original average displacements and standard deviations signals a diverse set of characteristics within the training data. Post-stabilization, although there is a noticeable reduction in these values, high maximum values hint at outliers or instances where stabilization is less effective. The Validation Dataset shows higher mean values in both original and stabilized displacements than the Train Dataset, suggesting that it might contain more challenging data for stabilization. This is reinforced by the high maximum values in both original and stabilized data, indicating the presence of extreme cases that are less responsive to stabilization. The Test Dataset, with the lowest variability in the original data, demonstrates the effectiveness of stabilization. However, the closeness of mean values for stabilized and original data suggests a different data characteristic or stabilization behaviour than the other datasets.

Effectiveness and Limitations of Stabilization

The effectiveness of the stabilization process across all datasets is evident (Table. 4.3), yet the high maximum values in the stabilized data, particularly in the Train and Validation datasets, point to scenarios where stabilization is less effective. These cases could be attributed to specific data conditions or characteristics resistant to the stabilization technique applied. Understanding these limitations is crucial for further refining the stabilization process.

The stabilization process, as applied to the training, validation, and Test datasets, predominantly showcases a significant improvement in stability. This is evidenced by the high number

Statistic	Dataset	Orig. Avg Disp.	Stab. Avg Disp.	Orig. Std Dev.	Stab. Std Dev.
Count	Train	47	47	47	47
	Val	9	9	9	9
	Test	12	12	12	12
Mean	Train	28.32	8.90	42.69	8.97
	Val	46.41	15.21	64.90	25.28
	Test	10.45	8.91	15.60	16.43
Std Dev	Train	63.23	33.85	96.65	25.97
	Val	47.99	34.52	56.85	51.34
	Test	9.31	26.17	20.80	41.25
Min	Train	1.29	0.23	1.01	0.05
	Val	2.27	0.16	1.66	0.13
	Test	2.54	0.15	1.60	0.11
Max	Train	344.06	198.02	506.31	113.76
	Val	146.06	103.00	136.21	137.65
	Test	33.61	91.35	59.08	136.52

Table 4.3: ***Comparative Analysis of Stabilization Metrics Across Train, Validation, and Test Datasets***
This table summarises key statistics, including the count, mean, standard deviation, minimum, and maximum values for original and stabilized average displacements and standard deviations. It effectively illustrates the variations and effectiveness of the stabilization process across different datasets.

of instances classified under 'Highly Significant Stability' across all datasets. Specifically, the Train dataset exhibits the highest number of such instances (42), followed by the Test (10) and Validation (8) datasets. This consistency in results indicates the general effectiveness of the stabilization technique in reducing the average displacement significantly, a key goal in many data processing and analysis tasks.

Notably, no instances in any of the datasets fall into the 'Equal Stability' category. This absence suggests that the stabilization process consistently alters the displacement values, indicating that the process is not neutral. In other words, the stabilization technique consistently improves or degrades the stability without maintaining the original state.

Moreover, there are a few instances where the stabilization results in 'Higher Post-Stabilization' displacement. Specifically, the Train dataset has 9 instances, whereas the Validation and Test datasets have 1 and 2 instances, respectively. While these numbers are relatively low compared to the instances of significant stability improvement, they are crucial for understanding the limitations of the current stabilization method. These instances might indicate specific scenarios or data characteristics where the stabilization process is less effective or counterproductive. However, the standard deviations for these differences are relatively high, indicating variability in the stabilization's effectiveness across different conditions. This variability is attributed to variations in image quality, which results in under-segmentation and, therefore, shifts the centroid of the mask contours.

Qualitatively, the stabilized images exhibited marked improvements in the uniformity of anther positioning. In time-lapse sequences, post-stabilization images showed anthers consistently positioned at the frame's centre, contrasting starkly with the pre-stabilization images. These results are visually depicted in Fig. 4.3, illustrating the before-and-after scenarios of anther

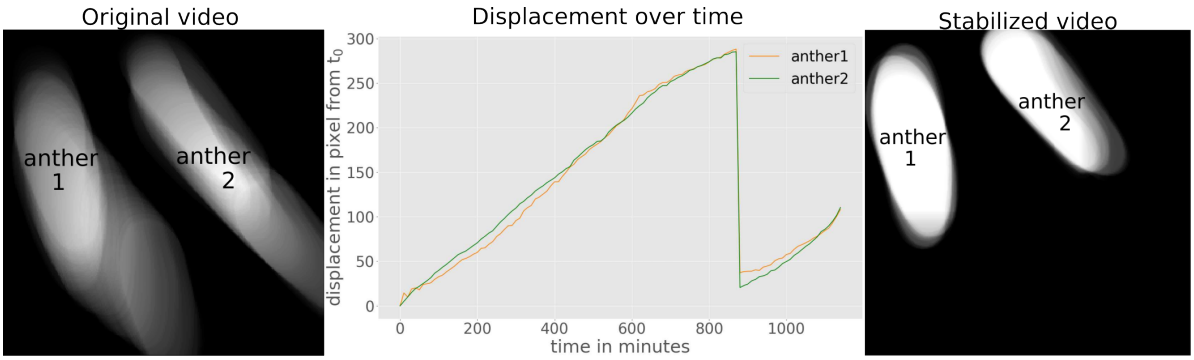

Figure 4.3: ***Motion compensation of anthers.***
The above figure shows two aspects. 1. tracking the anthers by the generated masks, 2. stabilising anther movement by an affine transformation. In one extreme case, the anthers exhibit considerable movement over time. Moreover, the gradual plant growth will result in a monotonic shift of the anther sample over time and require human intervention to re-position the plant so it does not leave the imaging focal plane. The monotonic movement over time until ≈ 800 frames and a manual reset after that.

stabilization.

While the stabilization model performed effectively in most cases, it exhibited limitations in scenarios involving rapid or extensive anther movements between frames. In such cases, the translation-only model was inadequate to correct the anther position, resulting in minor discrepancies. However, these were relatively rare occurrences involving motion beyond typical in our dataset.

Compared to more complex stabilization techniques, such as affine or elastic transformation models, our translation-only model offered a more computationally efficient solution with sufficient accuracy for most standard applications. The model's simplicity allowed for faster processing times, averaging ≈ 300 fps, which is particularly advantageous for high-throughput image analysis.

4.1.2 Conclusions and future directions

Integrating the U^2-Net model for anther segmentation and a translation-only model for anther stabilization represents a significant advancement in high-throughput image analysis within plant biology. While the approach demonstrated substantial effectiveness and efficiency, ar-

Category	Train	Validation	Test
Highly Significant Stability	42	8	10
Equal Stability	0	0	0
Higher Post-Stabilization	9	1	2

Table 4.4: ***Categorized Displacement Variations Across Train, Validation, and Test Datasets.***
The table illustrates the effectiveness of the stabilization process, highlighting the predominant occurrence of significant stability improvements and instances of increased displacement post-stabilization.

Description	Videos	Frames	Anther datapoints
Wild-type control	11	1071	2058
TCX	15	2437	4857
Control(21°)	30	5111	9803
Heat treated(30°)	43	5739	10,706
Heat treated(34°)	5	770	1540
Tetraploid(wild type)	24	2373	4375
Tetraploid(atm)	3	540	720

Table 4.5: ***Summary of unseen dataset composition for anther segmentation analysis.***
The table categorizes the data into experimental groups, including control and treatment conditions. It details the number of videos and frames analyzed for each group and the total count of anther datapoints extracted from the frames. This stratification ensures a comprehensive representation across different experimental conditions

eas for further refinement have been identified. These include enhancing the model's ability to handle complex anther movements and improving segmentation accuracy in low-contrast scenarios.

Our methodology effectively reduced the positional variance of anthers through rotational and translational adjustment and has been crucial in ensuring consistent anther positioning and orientation. This has been particularly beneficial for accurate morphological analysis in time-lapse studies. However, the challenges faced in scenarios with rapid or extensive morphological changes highlight the need for more advanced techniques.

As the next step in our research, image registration offers a solution to align images with high precision, facilitating detailed temporal analysis of anther development under varying conditions. This technique's integration would enable precise images overlaying across different times and conditions, enhancing our capability to track and analyze morphological changes, where the implementation of StyleGAN3 presents an exciting avenue. By utilizing StyleGAN3's capabilities in generating realistic images, we can generate a continuum of anther morphologies, thus enabling a deeper understanding of phenotypic variations under different conditions.

4.2 Results of Object Localization in Meiocyte Detection

In our study, we utilized EfficientDet models (D0, D1, D2) and their ensemble for the localization of meiocytes in high-resolution biological images. The ensemble approach's performance is particularly noteworthy in the test dataset, as detailed in Table 4.6. The ensemble model achieved a precision of 0.7732, indicating a high degree of accuracy in identifying meiocyte instances correctly. The recall rate for the ensemble model in the test set is 0.7853, a crucial metric in biological image analysis. This high recall rate ensures that the model effectively detects the majority of relevant meiocytes, a critical factor in avoiding missing key biological data. Moreover, the ensemble model exhibits a balanced performance in the test set, with an F1 score of 0.7779. This score is essential for evaluating the model's effectiveness in balancing precision and minimizing false negatives. Such a balanced approach is vital in our field, where detecting as many meiocytes as possible without compromising detection accuracy is crucial.

Metric	Validation Set				Test Set			
	D0	D1	D2	Ensembled	D0	D1	D2	Ensembled
Precision	0.9659	0.9750	0.9767	0.9626	0.7903	0.7784	0.7867	0.7732
Recall	0.7685	0.9106	0.9135	0.9233	0.7641	0.7665	0.7757	0.7853
F1 Score	0.8473	0.9385	0.9432	0.9409	0.7758	0.7710	0.7779	0.7779
IoU overlap	0.7518	0.7845	0.7849	0.7755	0.8326	0.8097	0.8234	0.8214

Table 4.6: *Object Detection Performance Metrics*
Metrics calculated with an IoU Threshold of 0.5 for the EfficientDet models (D0, D1, and D2) and the Ensembled model on both validation and test sets. The table includes precision, recall, F1 score and IoU overlap metrics.

Additionally, the Intersection over Union (IoU) score for the ensemble model in the test set is 0.8214. This high IoU score indicates that the model's localized bounding boxes are closely aligned with the actual boundaries of the meiocytes, which is vital for accurate morphological analysis and quantification. The ensemble performance across these key metrics - precision, recall, F1 score, and IoU - is comprehensively presented in Table 4.6. This table provides an in-depth analysis of the ensemble's capabilities and highlights its efficiency in handling the complexities inherent in biological image analysis.

Fig. 4.4 shows meiocyte localization across different frames. Each frame in the image highlights specific properties of meiocyte localization, such as the position and number of meiocytes within each frame. The visualization likely demonstrates the accuracy and efficiency of the object detection models in identifying and localizing meiocytes in various conditions. The differences in the frames may illustrate variations in meiocyte density, distribution, or other frame-specific attributes, providing a comprehensive view of the model's performance across diverse scenarios.

The visualization of meiocyte localization across different frames corroborates the statistical results from the ensemble of EfficientDet models, demonstrating their significant implications in our field. The visual consistency in meiocyte detection aligns with the high precision rate of the model, confirming that the identified meiocytes are true positives - a critical aspect for avoiding data misinterpretation and ensuring reliability in analyses. The comprehensive detection capability, evident in the high recall rate, minimizes the risk of omitting vital biological information. The F1 and IoU scores, observable through the accurate and consistent localization across frames, further reinforce the model's overall efficiency.

4.2.1 COCO Evaluation

In our study, the application of the COCO (Common Objects in Context) evaluation framework was critical for assessing the performance of object localization models, particularly the EfficientDet models (D0, D1, D2) and their ensemble, in localizing meiocytes in high-resolution biological images. The COCO framework, recognized for its comprehensive and rigorous evaluation metrics, provides an essential benchmark for evaluating these models. The results, presented in Table 4.7, encompass various metrics such as Average Precision (AP) and Average Recall (AR), offering a detailed analysis of model performance.

The AP values for the ensemble model in the test set, particularly at the AP @0.50 threshold, reached 0.968, indicating high precision in the model's positive predictions. This high AP @0.50 demonstrates the model's exceptional accuracy in detecting meiocytes with a 50% IoU

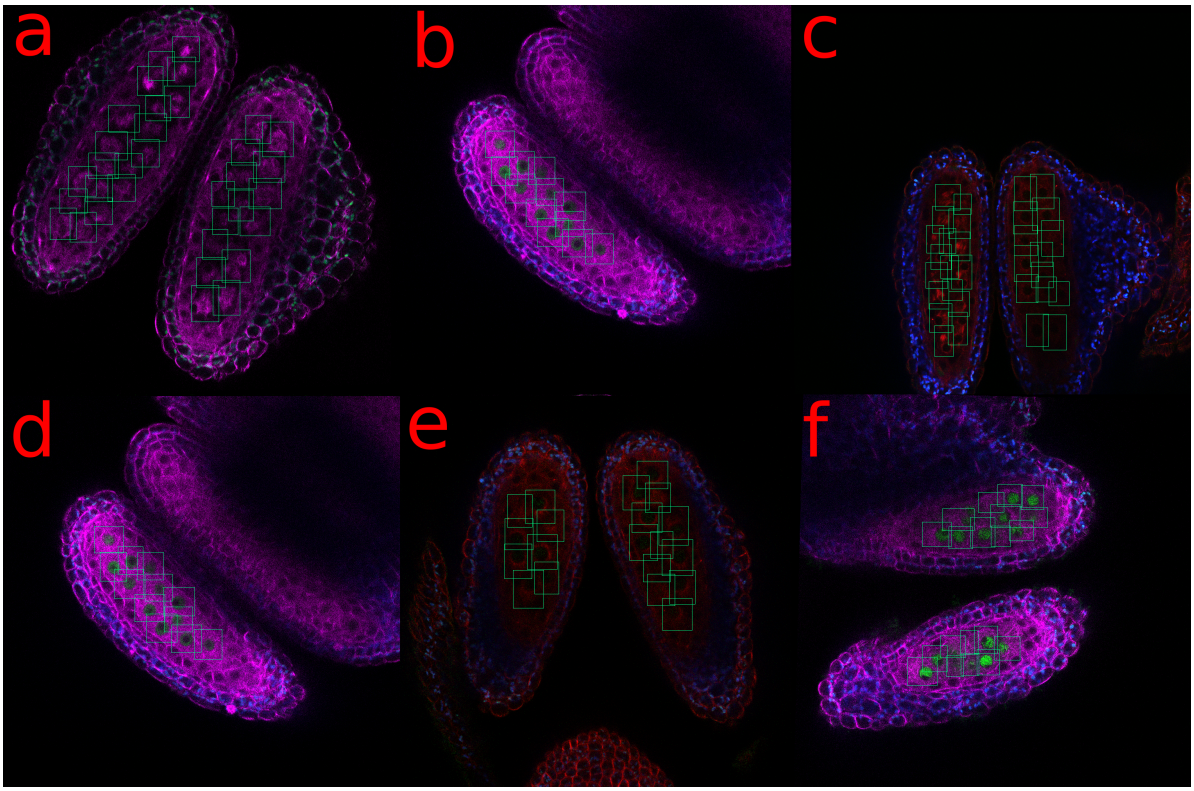

Figure 4.4: ***Meiocyte localization in anthers.***
Meiocyte localization in six distinct frames (a-f), demonstrating the EfficientDet ensemble model's consistent detection across varying densities and imaging conditions, validated by precision and recall rates, F1, and IoU scores.

threshold. Furthermore, the ensemble model's AP @0.75 in the test set is 0.754, suggesting a reliable performance even at a more stringent IoU threshold.

Average Recall (AR) metrics further elucidate the model's capabilities. For instance, in the test set, the ensemble model's AR @0.50 : 0.95 is 0.037, and its AR @0.50 is 0.333. These values indicate the model's effectiveness in detecting relevant instances across various IoU thresholds. Notably, the AR @0.75 for the ensemble model in the test set, which stands at 0.678, reveals its proficiency in detecting meiocytes with higher overlap accuracy.

The COCO evaluation framework's significance extends to providing a standardized benchmark against industry standards, thereby lending reliability and credibility to our results. This standardization is indispensable for objective model comparison within the machine learning and biological research communities. Moreover, the detailed analysis of AP and AR across different thresholds and object sizes is invaluable for understanding each model's performance in detecting objects of varying scales. The differentiation in performance metrics across small, medium, and large object sizes, as shown in the table, is particularly insightful. For instance, the AP-Large and AR-Large scores for the ensemble model in the test set are 0.758 and 0.793, respectively, highlighting its adeptness in detecting larger meiocytes with high precision and recall.

Utilizing the COCO evaluation method is integral for benchmarking against established standards and identifying specific strengths and weaknesses of each model. This is critical for targeted improvements and optimizations in future iterations of the models. The detailed quan-

tification provided by the COCO metrics enables a robust comparison of models, particularly highlighting the advantages of the ensemble approach in our study. The ensemble model consistently shows higher AP and AR scores across various thresholds and object sizes in the test set. This underscores its superior performance in accurately detecting and localizing meiocytes under diverse conditions.

Metric	Validation Set				Test Set			
	D0	D1	D2	Ensemble	D0	D1	D2	Ensemble
AP @0.50:0.95	0.335	0.477	0.465	0.470	0.627	0.576	0.543	0.625
AP @0.50	0.703	0.856	0.887	0.873	0.940	0.947	0.951	0.968
AP @0.75	0.253	0.481	0.521	0.455	0.775	0.667	0.623	0.754
AP-Small	0.000	0.000	0.000	0.000	-1.00	-1.00	-1.00	-1.000
AP-Medium	0.318	0.457	0.467	0.451	0.621	0.570	0.534	0.620
AP-Large	0.669	0.730	0.738	0.741	0.744	0.719	0.679	0.758
AR @0.50:0.95	0.395	0.535	0.613	0.535	0.678	0.633	0.589	0.678
AR @0.50	0.213	0.243	0.283	0.239	0.339	0.313	0.278	0.333
AR @0.75	0.033	0.035	0.039	0.033	0.037	0.035	0.039	0.037
AR-Small	0.000	0.000	0.000	0.000	-1.00	-1.00	-1.00	-1.00
AR-Medium	0.378	0.523	0.529	0.520	0.674	0.629	0.589	0.673
AR-Large	0.715	0.769	0.788	0.788	0.785	0.751	0.726	0.793

Table 4.7: *COCO Evaluation Results*
Validation and Test Sets for EfficientDet0, EfficientDet1, EfficientDet2, and the Ensembled Model. AP: Average Precision, AR: Average Recall. 'Small', 'Medium', and 'Large' refer to the area of detected objects.

4.2.2 Ensemble approach and non-maximum suppression

Furthermore, in our study, an ensemble approach was utilized to leverage the collective strengths of the EfficientDet models (D0, D1, D2). This methodology significantly enhanced the detection accuracy of meiocytes within the densely structured anthers. A key component in this enhancement was the application of non-maximum suppression (NMS). NMS played a critical role in minimizing redundant bounding boxes, thereby improving the precision of meiocyte identification.

The effectiveness of this approach is evident when examining the number of detections across both the validation and test sets, as detailed in Table 4.8. For instance, in the validation set, the ensemble method resulted in 3125 detections, compared to 2493, 2298, and 2457 detections for EffDet0, EffDet1, and EffDet2, respectively. In the test set, the ensemble approach led to 9041 detections, demonstrating a higher detection count than any individual model (EffDet0: 8585, EffDet1: 8779, EffDet2: 8876). This increase in detections by the ensemble method indicates its improved capability in identifying meiocytes.

Several post-processing steps were integrated to refine the localization results further. Detections outside the U^2-Net segmentation masks were excluded to maintain focus on relevant anther regions. A size-based exclusion criterion was employed, where detections exceeding a predetermined threshold were discarded. This criterion, informed by our understanding of meiocyte dimensions, was instrumental in eliminating detections that were unlikely to be actual meiocytes.

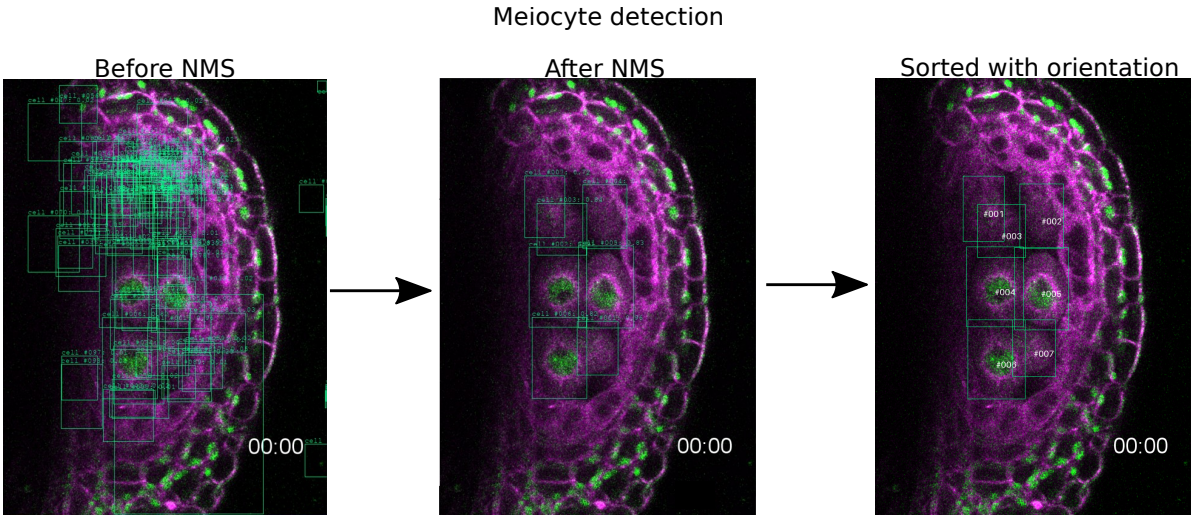

Figure 4.5: ***Comparative visualization of object detection on a single frame.*** The left image illustrates the preliminary detection phase with multiple bounding boxes per cell, indicating a high number of potential detections before applying Non-Maximum Suppression (NMS). After applying NMS, the right image displays the post-processing results, showcasing fewer bounding boxes, each corresponding to distinct cell localizations. This highlights NMS's effectiveness in refining the object detection output by eliminating redundant boxes and clarifying the visualization of individual cells. Timestamped at 00:00, the images capture the precise moment of analysis in the detection sequence.

	Validation Set				Test Set			
	D0	D1	D2	Ensembled	D0	D1	D2	Ensembled
Count	2493	2298	2457	3125	8585	8779	8876	9041

Table 4.8: Validation and Test Set Counts

Another of our approaches was the sorting and numbering of localized meiocytes based on their spatial positioning within each anther as shown in Fig. 4.6. This methodology enabled the tracking of individual meiocytes across different developmental stages. The practical utility of this technique is demonstrated in Fig. 4.5, where meiocytes within anthers are effectively numbered. This process facilitates detailed biological analysis and underscores the precision and utility of our ensemble approach in localizing and tracking meiocytes.

4.2.3 Limitations and future work

The localization section, underpinned by the data detailed in Table 4.9, has showcased the ensemble model's capability to process extensive datasets comprising various experimental groups, including control and treatment conditions. The model effectively analyzed thousands of frames and localized many meiocytes, demonstrating robustness and scalability.

Integrating a cell wall marker will refine the segmentation-localization merge, enhancing the model's accuracy in complex imaging scenarios. The potential shift towards models like YoLoV8 for real-time analysis, with our ensemble method currently operational at $\approx$ 12 frames per second, a promising indicator for future real-time applications. This groundwork paves the way for the seamless adoption of advanced deep learning architectures, ensuring readiness for the

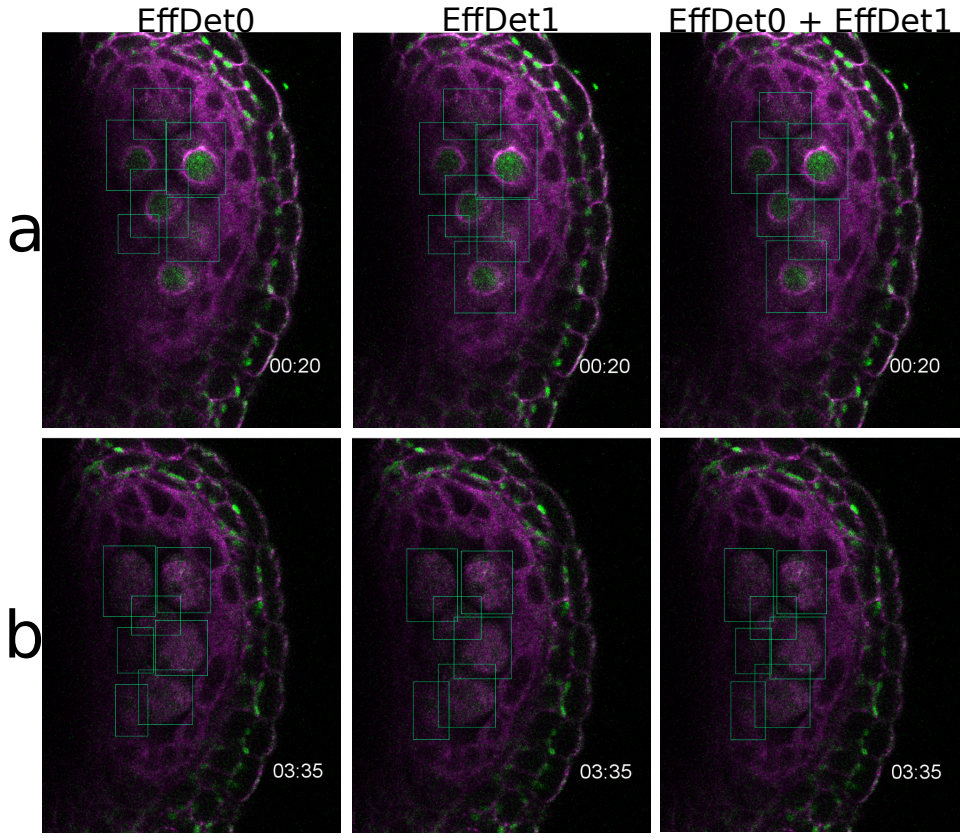

Figure 4.6: ***Ensembled approach - a comparison of EffDet models with ensembling.***
Sequential frames(a-b) depicting meiocyte localization using an ensemble of EfficientDet models. Each frame, labelled with time stamps at 00:20 and 03:35, shows the precision of bounding box placement over the meiocytes, illustrating the effectiveness of ensembling multiple network outputs.

Description	Videos	Frames	Anther count	Localized datapoint
Wild-type control	11	1071	2058	20409
TCX	15	2437	4857	49776
Control(21°)	30	5111	9803	83997
Heat treated(30°)	43	5739	10,706	60385
Heat treated(34°)	5	770	1540	6955
Tetraploid(wild type)	24	2373	4375	44847
Tetraploid(atm)	6	540	720	18112

Table 4.9: ***Summary of unseen dataset composition for anther and meiocyte localization analysis.***
The table categorizes the data into experimental groups, including control and treatment conditions. It details the number of videos, frames analyzed for each group, the total count of anther datapoints and localized meiocytes extracted from the frames. This stratification ensures a comprehensive representation across different experimental conditions and localization analysis.

increasing demand for speed without compromising precision in biological image analysis.

4.3 Meiocyte tracking in stabilized anthers

Building on the foundation of successful anther segmentation, stabilization, and meiocyte localization, we transitioned to the critical task of meiocyte tracking. Our tracking methodology, described in preceding chapters, involved novel linking algorithms designed to monitor and record the temporal progression of individual meiocytes through the stages of development.

4.3.1 Tracking accuracy and performance metrics

In our evaluation of the tracking system, we critically analyzed its performance using a comprehensive set of metrics, as summarized in Table. 4.10. The Multiple Object Tracking Accuracy (MOTA) and Multiple Object Tracking Precision (MOTP) were utilized alongside the Continuity Rate and Fragmentation Rate to provide a multidimensional assessment. The MOTA value of 0.89 for the training dataset indicates high accuracy in maintaining object identities due to missed tracks, false positives, or identity switches. The MOTP value of 0.87, reflecting the precision of object localization, corroborates this observation of high performance. The Continuity Rate is 0.9286, suggesting a strong ability of the system to track objects once they are detected consistently. However, the discrepancy between this rate and the MOTA implies that while tracking is consistent, it might occasionally start late (missing the initial frames) or involve some false positives. The Fragmentation Rate of 0.0 in the training set reinforces the system's strength in maintaining continuous tracking once an object is detected without losing it in subsequent frames.

In contrast, the validation and test datasets show perfect Continuity Rates and zero Fragmentation Rates, indicating no loss of tracked objects once identified. However, the MOTA values for these sets, 0.85 for validation and 0.87 for testing, again highlight inaccuracies due to initial misses or false positives. The higher MOTP values in these datasets (0.95 for validation and 0.98 for testing) suggest greater object localisation precision than the training set. This discrepancy across the datasets might be attributed to variances in the complexity of the scenes or the object densities, challenging the system differently in terms of initial object detection and identity maintenance.

Overall, these above results demonstrate the system's robustness in tracking continuity and precision in localization. However, they also highlight areas for improvement in initial object detection and reducing false positives, particularly in more complex scenarios like the high density of meiocyte in RSS (Restricted Search Space), contributing highly to object switching.

Dataset	MOTA	MOTP	Continuity Rate	Fragmentation Rate
Train	0.89	0.87	0.9286	0.0
Validation	0.85	0.95	1.0	0.0
Test	0.87	0.98	1.0	0.0

Table 4.10: ***Tracking metrics***
Average MOTA, MOTP, Continuity Rate, and Fragmentation Rate values for Train, Validation, and Test datasets.

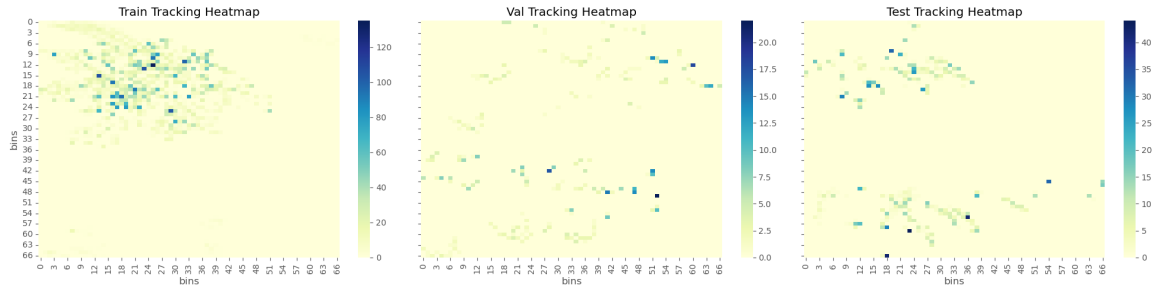

Figure 4.7: *Tracking heatmap*

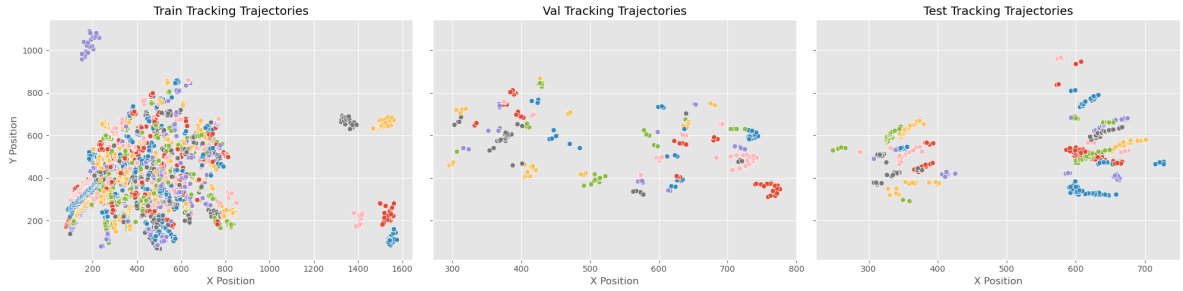

Figure 4.8: *Tracking trajectories*

4.3.2 Visualization of tracking results

The tracking visualizations for the 'Train', 'Val', and 'Test' datasets manifest distinct spatial and temporal patterns directly relevant to the previously discussed tracking metrics. The heatmaps in Fig. 4.7 correlate with the quantitative findings from the MOTA and MOTP scores, with the training dataset heatmap showing concentrated areas of high detection frequency. This aligns with the relatively lower MOTA score for the training set, potentially indicating localized instances of tracking inaccuracies or a greater complexity in the training environment that challenges the algorithm.

In contrast, the validation and test dataset heatmaps exhibit a more dispersed frequency distribution, aligning with the higher MOTA scores and suggesting a more uniform tracking accuracy across these datasets. The sparsity of high-intensity areas, particularly in the validation heatmap, reflects an environment with fewer tracking challenges or a more consistent application of the tracking algorithm.

The trajectory plots in Fig. 4.8 further substantiate these findings. In the training set, the dense convergence of trajectories could signify a range of meiocyte behaviours or an assortment of tracking scenarios, which could account for the observed variations in MOTA and MOTP scores. The less cluttered trajectories in the validation and test plots suggest a more straightforward tracking scenario, which could be a factor in the improved tracking precision denoted by the MOTP scores.

These visualizations depict the tracking algorithm's performance across different datasets. The heatmaps and trajectory plots serve as visual confirmations of the tracking continuity and fragmentation rates, with perfect continuity and zero fragmentation in the validation and test sets, suggesting a highly consistent tracking performance. Collectively, these visual representations provide a comprehensive overview of the tracking performance, offering both a macro and micro

Description	Videos	Frames	Anthers	Datapoint	Meiocytes
Wild-type control	11	1071	2058	20409	308
TCX	15	2437	4857	49776	420
Control(21°)	30	5111	9803	83997	373
Heat treated(30°)	43	5739	10,706	60385	734
Heat treated(34°)	5	770	1540	6955	117
Tetraploid(wild type)	24	2373	4375	44847	1774
Tetraploid(atm)	6	540	720	18112	155

Table 4.11: ***Dataset composition for meiocyte tracking analysis.***
The table categorizes the data into experimental groups, including control and treatment conditions. It details the number of videos, frames analyzed for each group, and the total count of other localized and tracked meiocytes extracted from the frames. This stratification ensures a comprehensive representation across different experimental conditions and localization analysis.

perspective that supports the algorithm's robustness as indicated by the tracking metrics.

4.3.3 Biological implications and future work

The tracking results provide valuable insights into the dynamics of meiocyte development during meiosis. The data is encapsulated in Table. 4.11 provides a quantitative backdrop for successfully tracking meiocytes across diverse experimental conditions.

Moreover, while effective, the present work's focus on two-dimensional tracking hints at the next evolutionary step in our tracking methodology: the exploration of three-dimensional stacked imaging. Adopting advanced generative models such as StyleGAN3, as hinted at in the segmentation section, will be pivotal. Such a development would address the challenges of cell switching in densely located cell populations, thereby enhancing the granularity and continuity of our tracking endeavours. This refinement is anticipated to bridge the gap between the high-throughput analysis showcased here and the nuanced complexities of biological systems, ensuring that the model remains at the forefront of innovation in meiocyte tracking technology.

4.4 Meiocyte classification in stabilized anthers

4.4.1 Overview of meiocyte classification results

Following the successful tracking of meiocytes, our study progressed to the critical phase of meiocyte classification. As detailed in earlier chapters, the classification methodology was designed to categorize meiocytes into distinct developmental stages based on morphological characteristics.

Parametric States	Validation Set		Test Set 1	
	Accuracy	F1 Score	Accuracy	F1 Score
Cell Shape	0.894	0.8979	0.9070	0.8977
Nucleus Position	0.895	0.8970	0.8864	0.8828
Nucleolus Position	0.9071	0.9070	0.8875	0.9023
Microtubular Array	0.835	0.8344	0.7916	0.8404
Chromatin (REC8)	0.8615	0.8558	0.8704	0.8515

Table 4.12: Performance metrics of classification network
Weighted categorical accuracy and F1 score for each parametric state

4.4.2 Classification accuracy and performance metrics

The critical evaluation of our classification model's performance centres on its accuracy and F1 score metrics, as tabulated for various parametric states. These metrics are pivotal in correctly assessing the model's ability to classify meiocytes into distinct developmental stages.

Cell Shape: The classification model's performance is quantitatively assessed using a confusion matrix(Fig. 4.9) and categorical accuracy scores(Table 4.12). In the confusion matrix, CS00 exhibits an accuracy of 76.4%, indicating model refinement and reduced ambiguity. This stage represents a critical juncture where the model differentiates between overlapping features of adjacent classes.

During the early prophase stage (CS01), the model's accuracy is lower, with an 89.0% correct classification rate, suggesting difficulties in capturing the onset of morphological changes. As the model progresses to CS02, covering stages from pachytene to meiosis II, accuracy improves to 87.8%, but misclassifications persist in transition states, highlighting the challenge of distinguishing subtle intra-state variations.

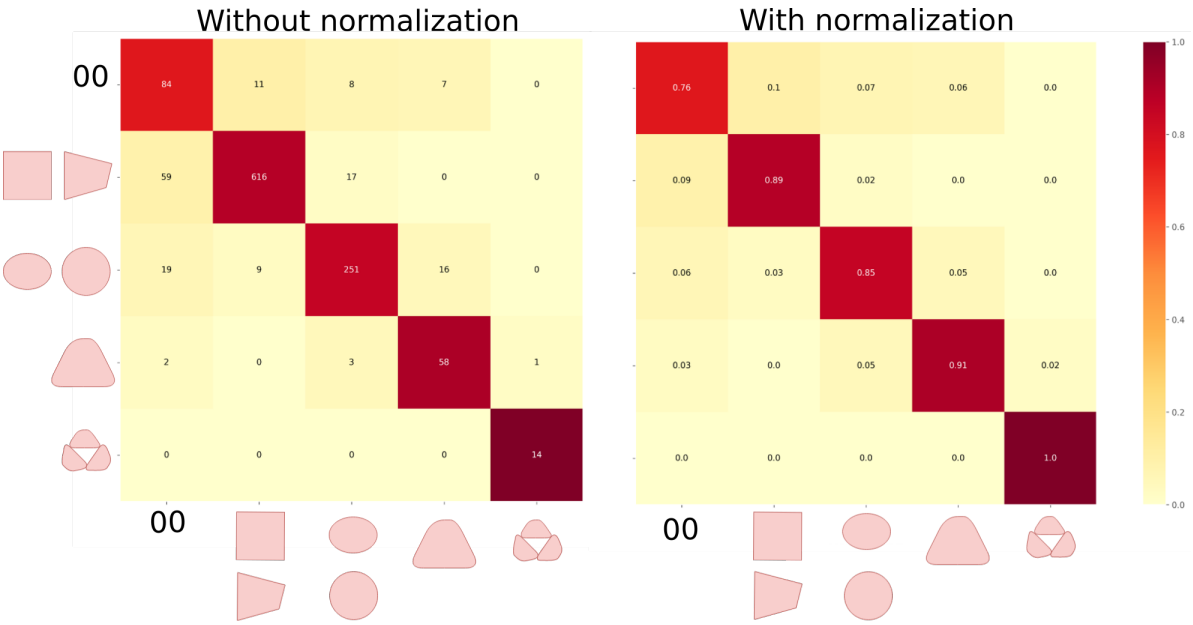

Figure 4.9: ***Confusion matrix for cell shape parameter***

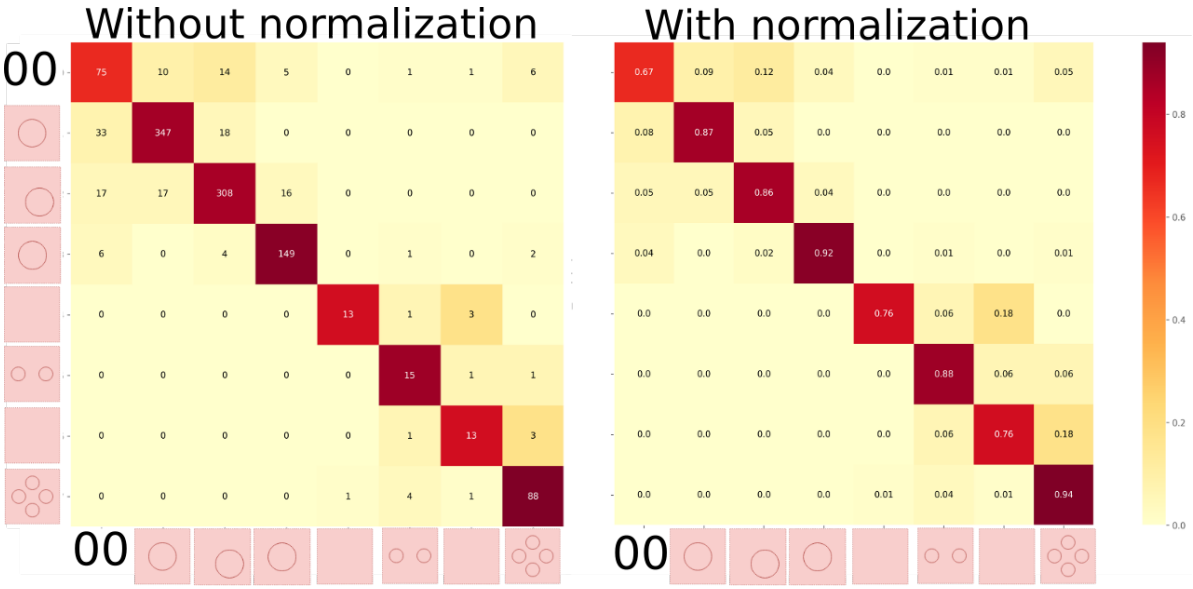

Figure 4.10: ***Confusion matrix for nucleus position parameter***

The pre-tetrad stage (CS03) shows a notable increase in model accuracy to 90.6%, indicating enhanced performance in recognizing distinct cell shapes. However, the presence of misclassifications, although fewer, emphasizes the potential for further refinement. The tetrad stage (CS04) stands out with 100% classification accuracy, showcasing the model's strength in detecting pronounced morphological transformations at this conclusive stage.

Nucleus Position: The classification of nucleus position (NP) is critical to understanding the stages of meiosis, and the confusion matrix provides a quantitative measure of our model's performance in this domain. Beginning with the ambiguous state NP00, our model demonstrates an accuracy rate of 76.5%, indicating strongly distinguishing confusing states. This ambiguity mirrors that seen in the cell shape classification, where early differentiation is paramount.

Progressing to NP01 and NP02, which depict early prophase and the transition to pachytene, the model exhibits improved accuracy, with positive rates of 89.7%and 87.6%, respectively. These stages are critical as they mark the commencement of visible morphological changes. The transition from NP02 to NP03, representing the diplotene to diakinesis state, shows the model's accuracy at 87.8%, reflecting its capability to discern the distinct chromosomal configurations characteristic of these stages.

As we move to NP04, indicative of the metaphase to anaphase transition where the nucleus is not visible, the model's performance aligns with the categorical accuracy and F1 score, with a high correct classification rate of 93.7%. This stage's clarity regarding nuclear absence may contribute to the model's higher accuracy.

For NP05, corresponding to telophase, where two nuclei are distinctly visible, the model achieves a classification rate of 88.2%, suggesting a strong ability to recognize the dual-nuclei structure that defines this stage. Similarly, NP06, which covers the second metaphase to anaphase transition with the nucleus again invisible, the model continues to perform well with a classification rate of 92.9%.

Our model excels at the tetrad stage, NP07, with an accuracy of 94%, showcasing its strength

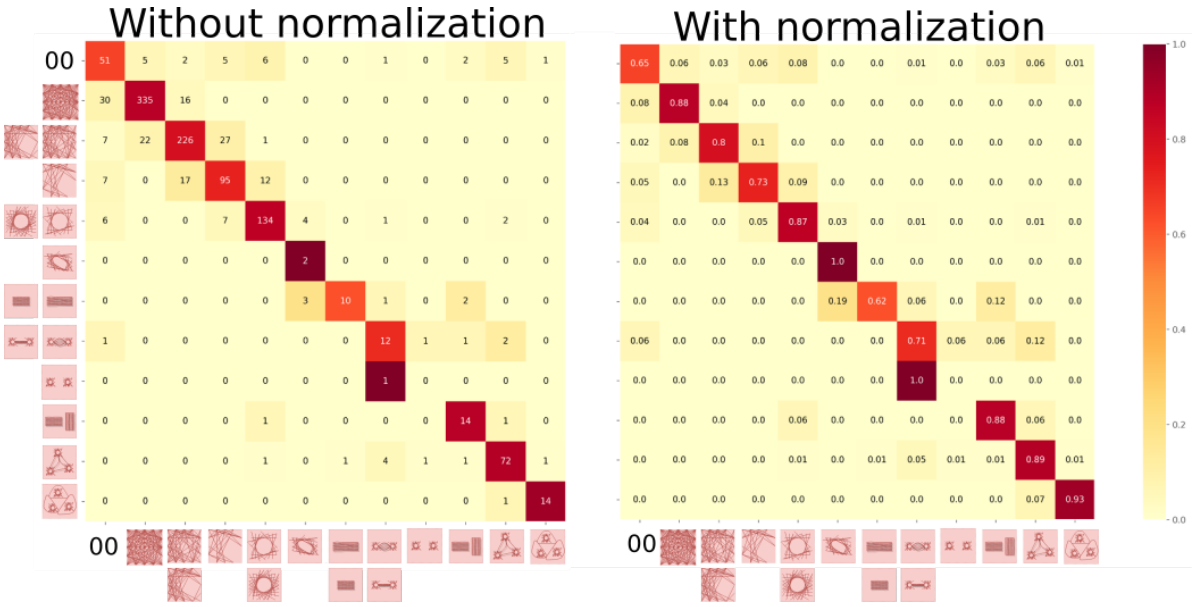

Figure 4.11: ***Confusion matrix for microtubular array parameter***

in identifying the culmination of meiotic processes. This high accuracy rate for the tetrad stage suggests the model is particularly adept at identifying the complex structural configurations present at this stage.

Microtubular Array: The microtubular (MT) array classification, integral to charting meiosis progression, reflects the complex interplay of cellular structures during meiotic phases. The provided image, though not directly accessible, likely showcases the transition of microtubule configurations from MT01 through MT11, representing distinct meiotic stages.

MT01 and MT02, indicative of the S/G2 phase through to zygotene, show increased model precision. The model identifies the S/G2 phase with an accuracy of 76.4% for MT01, improving to 81.4% for MT02. This improvement aligns with the visual progression in the image, where distinct microtubular formations become more apparent.

The model achieves an accuracy of 87.8% for MT03, corresponding to the pachytene stage, distinguished by specific chromosomal alignments and microtubule organization. The diplotene and diakinesis phase (MT04) presents a classification accuracy of 82.6%, where the confusion matrix indicates occasional misclassifications with adjacent stages.

The stages from MT05 through MT07, depicting nuclear envelope breakdown and transition to telophase I and interkinesis, show varying classification accuracy, from 72.3% for MT05 to 76.9% for MT07. These stages, characterized by significant cellular reorganization, are where the model's predictive capability is tested, as highlighted by the confusion matrix.

For the later stages, MT08 to MT10, covering meiosis II, the model's classification accuracy reaches its zenith, with an impressive 93.3% for MT10, indicative of its adeptness at identifying the later stages of meiosis II. The tetrad stage, MT11, is classified with 93.3% accuracy, underscoring the model's efficiency in recognizing the culmination of meiosis as tetrads are formed.

Nucleolus Position: The classification of nucleolus position (NoP) throughout meiosis offers

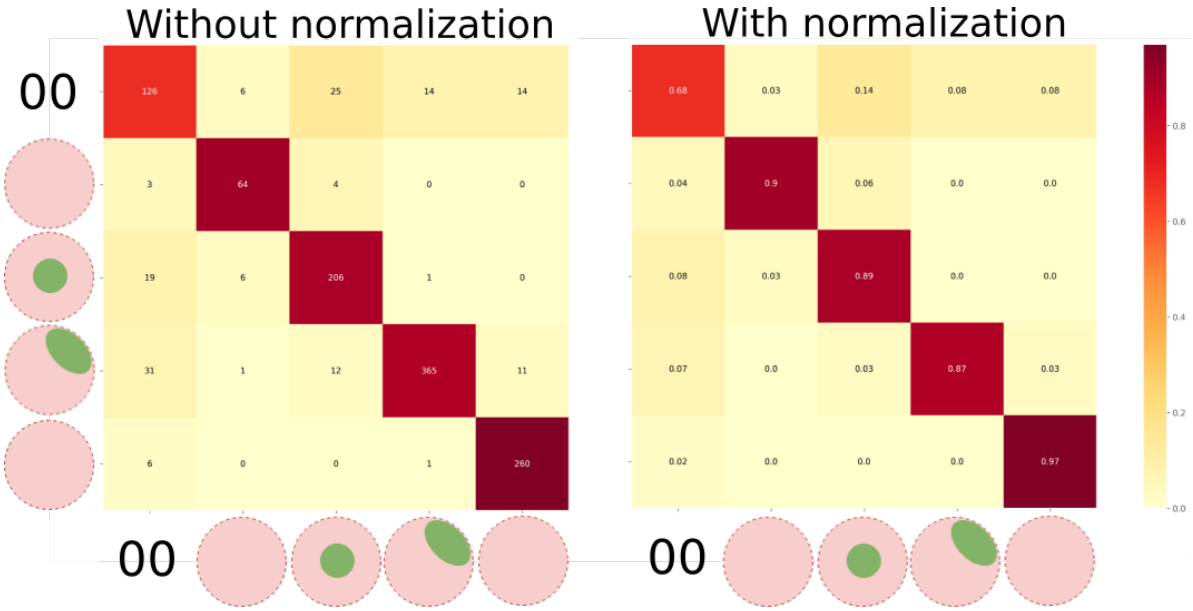

Figure 4.12: ***Confusion matrix for nucleolus position parameter***

the least complicated stages of cellular development. The confusion matrix for NoP classification details the model's ability to track the nucleolus through its dynamic changes.

For NoP01, where the nucleolus is absent during the S/G2 phase, the confusion matrix indicates a strong model performance with a high correct classification rate of 82.8%. However, there is some confusion with later stages, possibly due to variations in nuclear visibility that the model misinterprets.

The early leptotene stage, NoP02, characterized by a central nucleolus, is accurately classified with 89.4%, as the distinctive central positioning provides a clear signal for the model to detect. However, 6 misclassifications with NoP01 suggest that when the nucleolus begins to appear, it may be occasionally confused with its absence.

NoP03, marking the nucleolus's movement to the side of the nucleus visible until pachytene, presents a classification accuracy of 90.3%. This stage's classification is relatively reliable, reflected in the F1 score of 90.70%, indicating a strong balance between precision and recall.

The final stage, NoP04, where the nucleolus is no longer visible for the rest of meiosis progression, shows an exceptionally high accuracy of 95.6%. This suggests that the model adeptly identifies the absence of the nucleolus, a significant morphological marker.

Chromatin (REC8): The classification of chromatin dynamics, as marked by REC8 protein localization, traces the intricate process of meiotic division, with the confusion matrix offering quantitative insights into the model's classification accuracy across stages. Starting from the ambiguous state RC00, the model distinguishes this phase with an accuracy of 83%

As the meiotic division progresses, the model's ability to differentiate between stages becomes evident. For instance, in the S/G2 phase represented by RC01 and the early leptotene stage denoted as RC02, the model achieves accuracy rates of 89.4% and 90.3%, respectively. This suggests a high level of model precision in capturing the initial chromosomal dynamics marked by the REC8 localization.

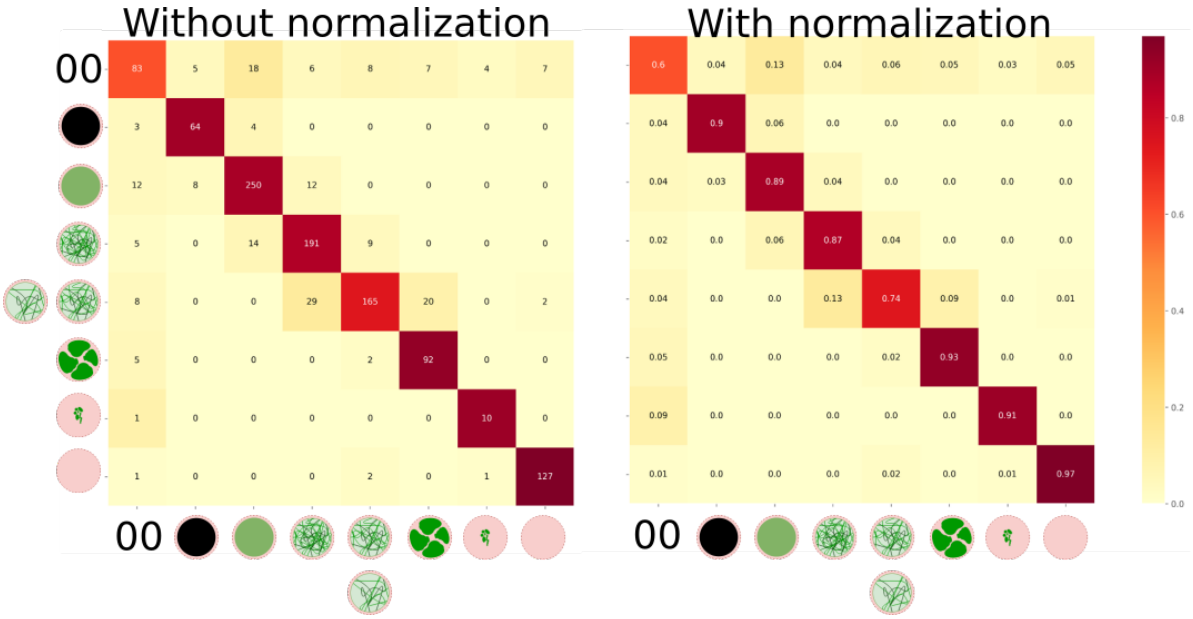

Figure 4.13: ***Confusion matrix for chromatin parameter***

Moving into the zygotene stages encapsulated by RC03, the model maintains a high accuracy of 87.8%, indicative of its capability to track the continuity of REC8 localization patterns. For the pachytene stage, RC04, the model's accuracy is 82.6%, signifying a reliable recognition of the unique chromosomal configurations.

RC05, unique to the diplotene diakinesis of prophase 1, exhibits an accuracy of 93.7%, reflecting the model's adeptness at identifying the distinct separation of homologous chromosomes. The transition to RC06, signifying the entry into metaphase-1 and visible until the onset of anaphase-1, shows the model's accuracy at 88.2%, demonstrating its efficiency in discerning the metaphase alignment of chromosomes.

Finally, RC07, where chromatin is not visible for the remainder of meiosis, is accurately classified with an impressive rate of 94%, underscoring the model's effectiveness in recognizing the completion of chromosomal segregation.

4.4.3 Biological implication and future direction

Implementing automated classification within the aMP framework has achieved a dual advantage: it has significantly accelerated the analysis process and eliminated the subjectivity inherent in human-based evaluation. This two-pronged enhancement is pivotal for accurately staging meiotic progression and facilitates the simultaneous assessment of interdependent meiotic stages. Such analysis is critical for identifying anomalies or potentially novel states that may emerge in mutant phenotypes or across different plant models.

Moreover, the classification system establishes an essential foundation for generating artificial cell variation models. By employing techniques such as transfer learning and variational autoencoders, we can simulate the appearance of cells under a spectrum of biological conditions. This capability is transformative, offering predictive insights into cell morphology and behaviour before the execution of actual imaging tasks. The long-term impact of this advancement is

profound, promising substantial strides in our ability to forecast cellular dynamics based on underlying biological parameters.

4.5 Evaluation of meiosis landmarks

The landmark and meiosis timeline evaluation dissects the analysis into three distinct segments, each addressing a crucial aspect of meiotic progression in *Arabidopsis thaliana*. Firstly, we scrutinize the impact of regrouping parametric states on the resolution of meiotic landmarks, a process that refines the delineation of critical meiotic transitions. Secondly, we draw correlations with previously published results to anchor our findings within the established body of knowledge and validate the robustness of our approach. The third segment is dedicated to achieving the core objective of this project: analysing the meiosis progression timeline in tetraploid genotypes, a domain previously unexplored due to the inherent complexities presented by polyploidy. Finally, we extend our analytical framework to include the *tcx5;6* mutants, broadening our investigation to encompass genetic variations that may offer new insights into the meiotic mechanisms of this model organism.

4.5.1 Effect of regrouping parametric states

The evaluation of meiosis progression landmarks is predicated on the analysis of thresholded z-normalized scores, approached in a piece-wise manner to delineate the transition times between successive meiotic landmarks, as described by Eq. 3.44. In contrast to the methodology reported by Prusicki et al. [3], which utilized ungrouped states for the calculation of landmarks, our project introduces a grouping paradigm as illustrated in Fig. 2.4. This initial exploration investigates the implications of this regrouping on the identification of meiotic states and the associated landmark scores defined by Eq. 3.43, offering a comparative analysis with the original findings, as depicted in Table 4.13.

Prusicki et al. [3] derived meiotic landmarks from ungrouped parametric state labels, highlighting the direct influence of these parametric classes on normalized scores and the definition of landmarks. The reclassification into grouped parametric classes, as evidenced in Table 4.13, elucidates a diminished resolution in prophase I, with the original landmarks $A_{0,\dots,12}$ being condensed into $A'_{0,\dots,9}$. Specifically, this regrouping results in the amalgamation of stages, leading to the loss of distinct stages for late leptotene, zygotene, and pachytene in prophase. At the same time, the characterization of meiosis-II states remains intact.

The regrouping effectively merges the distinct stages of late leptotene (A_2), zygotene (A_3 and A_4), and pachytene (A_5 and A_6), as these stages no longer exhibit individual z-scores in the regrouped analysis. For instance, the early leptotene stage originally marked with a z-score of 3.388 (A_1) is adjusted to 2.98 in the regrouped analysis (A'_1), indicating a subtle shift in the scoring methodology. Conversely, stages such as diplotene to diakinesis (A_7 to A'_4), metaphase-I to anaphase-I (A_8 to A'_5), and telophase-I to interkinesis (A_9 to A'_6) demonstrate a more pronounced alteration in z-scores, reflecting the regrouping's impact on the resolution of meiotic phase transitions. The preservation of meiosis-II states (A_{10} to A'_7 and A_{11} to A'_8) underscores the selective nature of the regrouping effect, with the tetrad state (A_{12} to A'_9) remaining unaltered, thus maintaining its critical role in the meiotic timeline.

Landmark	Z-score	Z-score regrouped	Landmark regrouped	State property
A_0	−0.18	−0.18	A'_0	S/G2 phase
A_1	3.388	2.98	A'_1	Early leptotene
A_2	1.14	–	–	Late leptotene
A_3	1.34	3.42	A'_2	Zygotene
A_4	3.20	–	–	Zygotene
A_5	3.02	2.92	A'_3	Pachytene
A_6	1.11	–	–	Pachytene
A_7	2.76	2.40	A'_4	Diplotene to Diakinesis
A_8	2.45	1.97	A'_5	Metaphase-I to Anaphase-I
A_9	2.63	2.02	A'_6	Telophase-I to Interkinesis
A_{10}	2.37	2.21	A'_7	Meiosis-II
A_{11}	2.23	2.23	A'_8	Meiosis-II
A_{12}	−0.45	−0.45	A'_9	Tetrad state

Table 4.13: *Effect of parametric grouping on landmark states and Z-scores*
The original landmarks [3] from the dataset-1 and their corresponding scores when analysed using regrouped classes lose the minute resolution in prophase-I, where $A_{2,3}$, $A_{4,5}$ and $A_{6,7}$ are combined to reduce the original number of landmarks by 3.

The refined analysis presented in this project delineates the consequences of regrouping parametric states on the resolution and identification of meiotic landmarks. By comparing these newly grouped landmarks against the original framework established by Prusicki et al. [3], this study underscores the refinement of parametric regrouping on the automated analysis of meiosis progression, offering a comprehensive overview of the methodological shifts and their implications for automated staging of meiocytes in meiosis landscape.

4.5.2 Analyzing meiosis progression across wild type dataset at 21°C

The aMP framework was applied to evaluate meiosis in a distinct wild type *Arabidopsis thaliana* control dataset, grown at 21°C by a different investigator, to ensure an unbiased analysis. This dataset, referenced as [32], was analyzed to calculate normalized z-scores, facilitating a comparative analysis with the primary dataset cultivated under similar conditions. A critical part of this analysis involved the comparison of the meiosis progression timelines. The aMP framework's capability to match and sometimes surpass manual analysis benchmarks was evaluated, particularly focusing on microtubular array dynamics that evolve throughout meiosis.

Comparing the progression timelines with the outcomes reported in [32], we adopted an offset measure to assess staged progression. This method involves disregarding data points where landmarks are only partially observable. Specifically, if landmark L_i signifies the progression from state A_{i-1} to A_i, it's mandatory that A_{i-2} and A_{i+1} is observed; otherwise, the piecewise timeline for that meiocyte isn't computed. Interestingly, the published analysis excludes the L_1 and L_9 transition entirely (Fig. 4.15), highlighting the importance of offset analysis in providing a more comprehensive understanding of meiosis progression.

The comparative analysis of the two datasets maintained at 21°C—designated as WT 21°C(MP) for the analysis conducted via the Automated Meiosis Progression (aMP) framework and WT

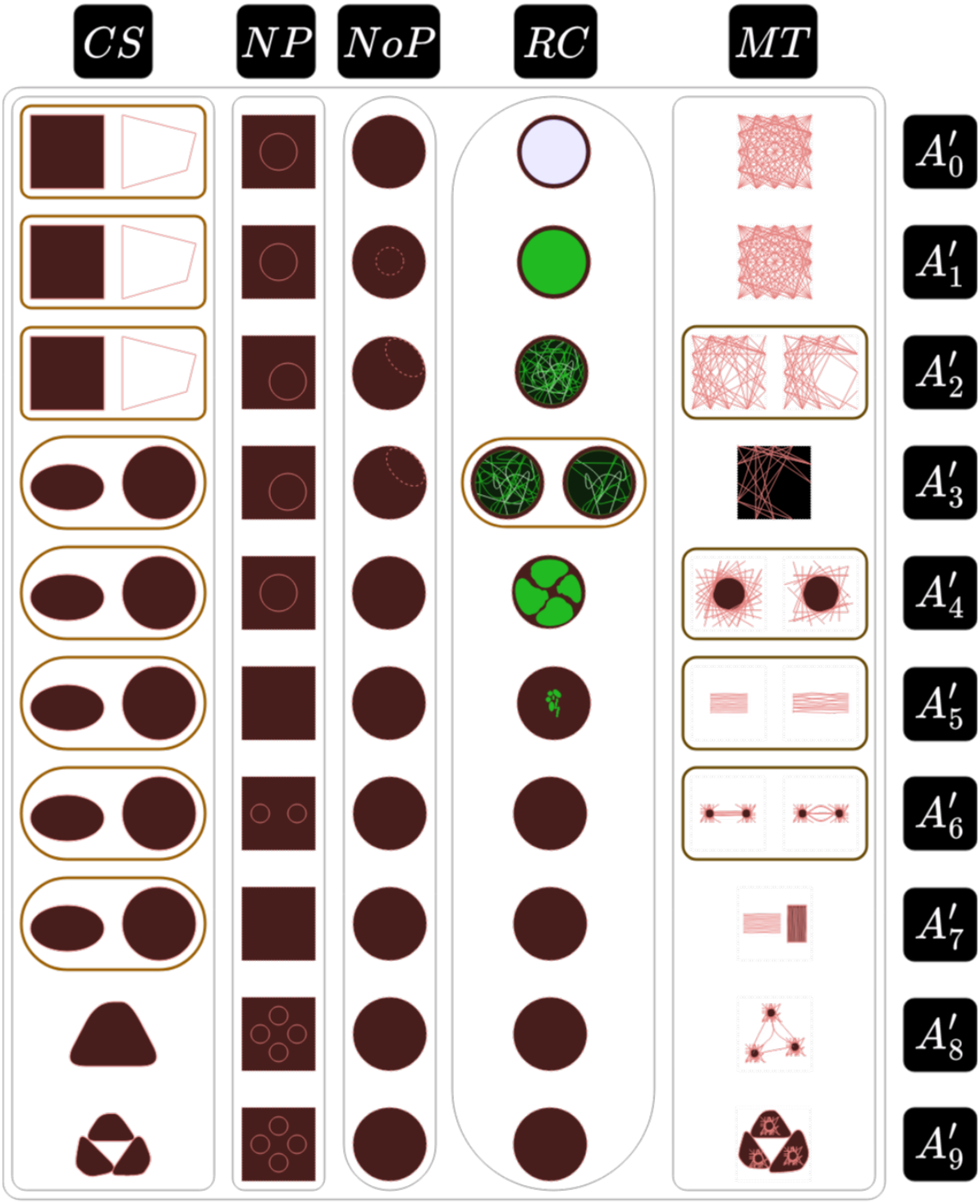

Figure 4.14: Regrouped landmark states and their appearance

Landmark states	WT 21°C(MP)	WT 21°C(JDJB)	State property
A'_0	−0.18	0.54	S/G2 phase
A'_1	2.98	3.24	Early Leptotene
A'_2	3.42	3.85	Early Zygotene
A'_3	2.92	3.42	Early Pachytene
A'_4	2.40	2.99	Diplotene to Diakinesis
A'_5	1.97	2.25	Metaphase-I to Anaphase-I
A'_6	2.02	1.95	Telophase-I to Interkinesis
A'_7	2.21	2.59	Meiosis-II
A'_8	2.23	2.54	Meiosis-II
A'_9	−0.45	0.26	Tetrad state

Table 4.14: Landmark score comparison between 2 wildtype A.thaliana dataset at 21°C

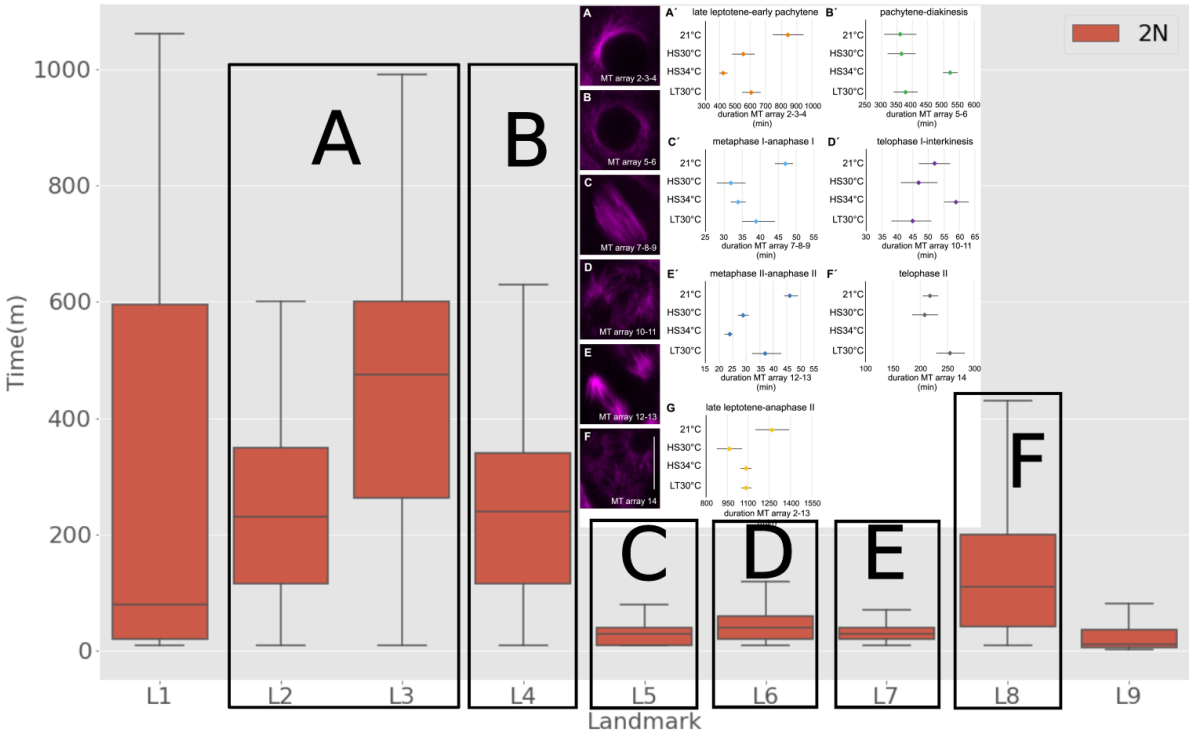

Figure 4.15: ***Sanity check for automated timeline progression.***
Automated timeline analysis of dataset-II compared with the manual analysis provided in [32] shows semblance in the piece-wise meiosis progression timeline.

21°C(JDJB) for the dataset evaluated in [32]—illuminates essential insights into the temporal dynamics of meiosis. Variations in z-scores across the spectrum of meiotic phases, from S/G2 to the tetrad state, highlight subtle yet significant differences that underscore a dataset-specific calibration of landmark identification while affirming the landmarks' inherent stability. This observation evidences the aMP framework's enhanced sensitivity and precision in charting the meiosis timeline, underscoring the critical importance of accurately defining the prophase timeline for an exhaustive understanding of meiosis progression.

For instance, the early leptotene stage, originally marked with a z-score of 2.98 for WT 21°C(MP), compared to 3.24 for WT 21°C(JDJB), underscores a slight calibration in landmark

identification between datasets. This trend of re-calibration is observed across various stages, including early zygotene and early pachytene, with corresponding adjustments in z-scores highlighting the framework's analytical consistency in featuring the exact stages in different constructed datasets.

Furthermore, the evaluation showcases a close approximation between automated and manual analysis timings for the progression through meiotic landmarks, as depicted in Table 4.15. Despite the automated approach introducing an offset for partially observed data points, the timelines closely align with manual estimates, underscoring the aMP framework's effectiveness. Such an alignment validates the automated approach and illustrates its potential to facilitate a generalized, user-adjustable framework for analyzing meiosis progression.

This detailed comparative analysis emphasizes the robustness and adaptability of the aMP framework in analyzing meiotic progression under varying conditions. By aligning closely with manually analyzed data while offering flexibility in data interpretation, a framework is a valuable tool for researchers studying meiosis in *Arabidopsis thaliana* and its potential variation based on the ploidy-ness of plants and particular gene insertion to check the difference in prophase-I timeline variation. The evaluation is restricted to the control dataset and did not include the heat shock experiments due to the lack of meiosis-II data points; also, we are using 12 original videos to analyse the result owing to its previously reported result, which is efficient for benchmarking.

4.5.3 Tetraploid timeline for meiosis progression

The comparative analysis of meiotic progression timelines between wild-type diploid (2N) and tetraploid (4N) *Arabidopsis thaliana* is illustrated in Figure 4.16 and Table 4.16. This analysis provides crucial quantitative insights into the temporal dynamics of meiosis across different ploidy levels.

In tetraploids, the onset of the S/G2 phase (A'_0) is characterized by a z-score of 1.39, starkly contrasting to the wild-type's -0.18. This suggests a protracted commencement or an extension of this phase. Such an extension may indicate a complex adaptive mechanism responding to the increased chromosomal content, necessitating a recalibration of cell cycle checkpoints to ensure genomic stability before meiotic entry.

Proceeding through the meiotic timeline, early leptotene (A'_1) and early zygotene (A'_2) stages in tetraploids manifest lower z-scores (1.80 and 3.64, respectively) relative to their diploid counterparts. This downshift in z-scores could reflect a decelerated progression, possibly due to

Landmark	Automated mean time	Automated mean time (offset)	Manual mean time
L2 + L3	765	791	845
L4	246	398	360
L5	33	50	47
L6	53	50	52
L7	35	48	46
L8	125	201	219

Table 4.15: Mean Time(m) in meiosis transition states

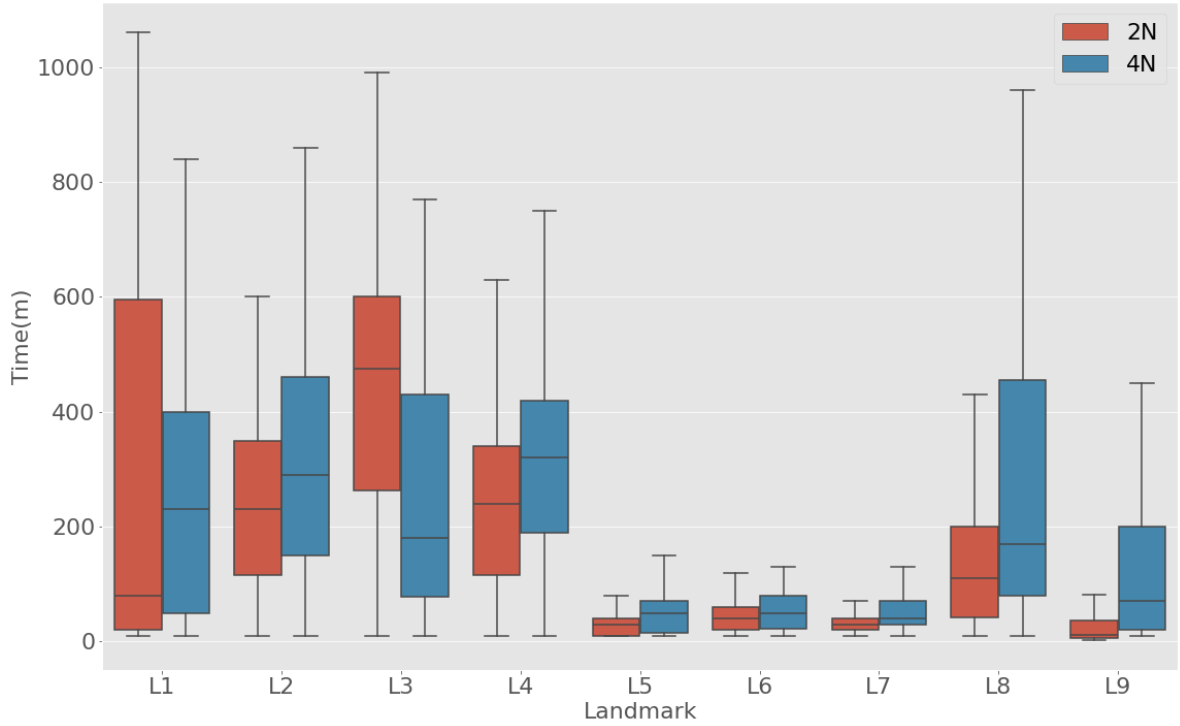

Figure 4.16: ***Comparing tetraploids to the wild type A.thaliana variants***

the augmented chromosomal interactions requiring meticulous homologous pairing and synapsis orchestration. This process is inherently more complex in tetraploids.

Notably, from diplotene to diakinesis (A'_4) onwards, the z-score convergence with the diploid wild-type at 2.99 implies that once tetraploids surpass the initial meiotic complexities, their subsequent progression is remarkably synchronous with that of the diploid. However, a subtle diminution in the z-score for the final meiosis-II phase (A'_8) from 2.23 in the diploid to 2.08 in the tetraploid suggests a slightly accelerated resolution of meiosis.

The extended early meiotic stages observed in tetraploid *A. thaliana* suggests a recombination landscape that differs from diploids, with implications for genetic diversity. In tetraploids,

Landmark states	WT (MP)	WT (JDJB)	TP (MP)	State property
A'_0	−0.18	0.54	1.39	S/G2 phase
A'_1	2.98	3.24	1.80	Early Leptotene
A'_2	2.99	3.85	3.64	Early Zygotene
A'_3	2.92	3.42	3.55	Early Pachytene
A'_4	2.40	2.99	2.99	Diplotene to Diakinesis
A'_5	1.97	2.25	2.46	Metaphase-I to Anaphase-I
A'_6	2.02	1.95	2.01	Telophase-I to Interkinesis
A'_7	2.21	2.59	2.43	Meiosis-II
A'_8	2.23	2.54	2.08	Meiosis-II
A'_9	−0.45	0.26	−0.14	Tetrad state

Table 4.16: The comparison of landmark states' Z-score between the 2 wild-type and tetraploid dataset.

the protracted duration of phases such as early leptotene and zygotene may facilitate a more thorough recombination process, enhancing genetic variability. Differences in meiotic timing between diploid and tetraploid *A. thaliana* reveal the impact of ploidy on meiotic control and cellular function. This ploidy-specific modulation of meiosis is likely an evolutionary strategy for handling complex chromosome segregation challenges inherent to polyploidy.

ATM gene insertion effects on pachytene checkpoint in tetraploid *A. thaliana*

The ATM (Ataxia-telangiectasia mutated) gene is known to play a crucial role in the pachytene checkpoint during meiosis by assessing homologous recombination and DNA double-strand break (DSB) repair. In tetraploid *Arabidopsis thaliana*, the absence or mutation of the ATM gene (atm-4N) is posited to weaken these checkpoint mechanisms, as ATM's function as a protein kinase activated in response to DSBs is diminished or absent. This reduction in ATM activity is particularly impacted in tetraploids, where the increased chromosomal number intensifies the challenges of chromosomal pairing and segregation.

Without a functional ATM, the pachytene checkpoint's scrutiny is hypothesized to be less stringent, leading to a reduced arrest at this stage, potentially resulting in incomplete DSB repair and inaccurate synapsis. This abbreviated arrest would manifest in the meiosis timeline as a shortened leptotene to pachytene stages (L1-L3) duration, as shown in Figure 4.17. These stages, particularly pachytene, would exhibit a decreased duration in atm-4N tetraploids, reflecting the reduced ATM kinase activity and its diminished role in DSB surveillance and repair.

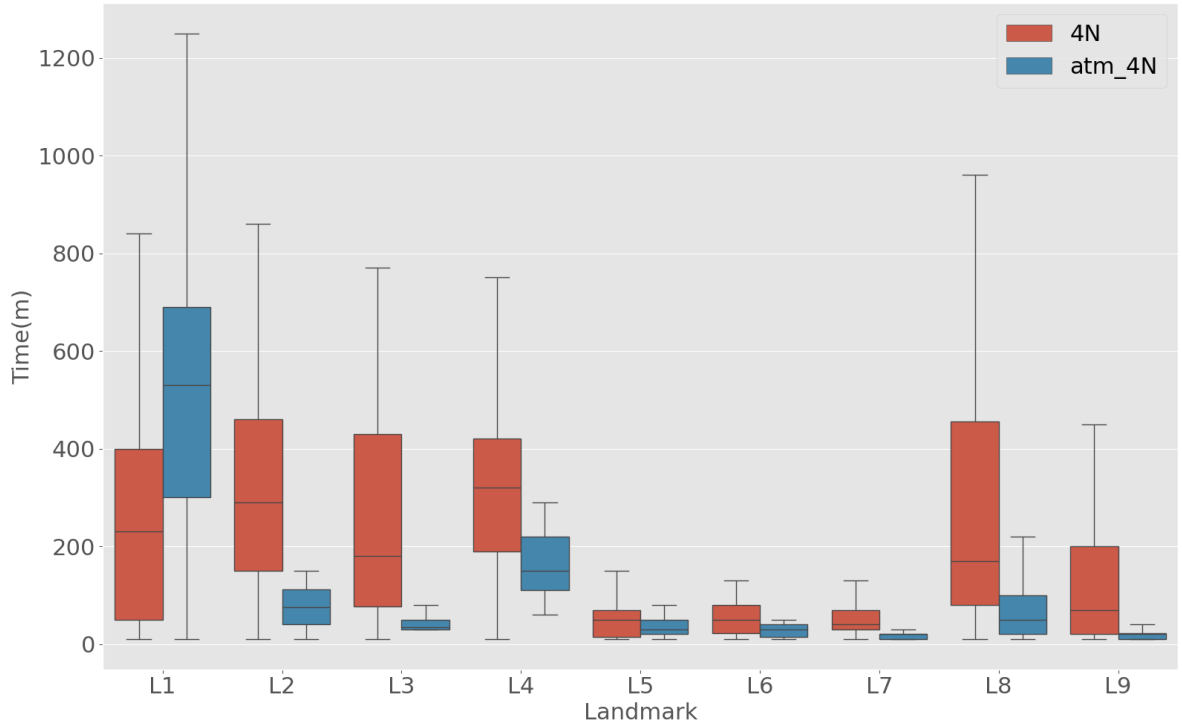

Figure 4.17: ***Comparing tetraploids to its atm-gene insertion A.thaliana variants***

In the post-pachytene stages (L4-L9), atm-4N tetraploids display heterogeneity in meiotic duration, with some phases potentially progressing faster than in wild-type tetraploids. This variability suggests that the weakened ATM-mediated pachytene checkpoint may allow cells to advance through the later stages of meiosis at an accelerated rate, even in unresolved DNA double-strand breaks (DSBs).

The dataset also indicates increased inter-stage temporal variability in atm-4N plants, implying less uniform progression through meiosis post-pachytene. This variability reflects the impaired role of the ATM gene in enforcing stringent DSB repair standards, leading to a more variable and potentially error-prone meiotic process. Consequently, the mutation of the ATM gene in tetraploid *A. thaliana* weakens the pachytene checkpoint, shortening the duration of early meiotic stages and allowing faster progression despite incomplete DSB repair.

4.5.4 Comparative Analysis of *tcx5;6* Mutants and Wild-Type Diploid *A. thaliana*

After characterizing the ATM gene effects in tetraploid *A. thaliana* with the Automated Meiosis Progression (aMP) framework, we now expand this method to explore *tcx5;6* mutants. The aMP framework's precise analysis of meiotic stages showcases its utility for wider applications, encompassing the evaluation of genetic variants in diploid species.

The *tcx5;6* mutants in diploid *Arabidopsis thaliana* represent a unique opportunity to examine the genetic control of meiotic progression, particularly about the TCX5 and TCX6 genes, which are believed to play roles in chromatin organization and gene expression during meiosis. Mutations in these genes may disrupt the tightly regulated sequence of meiotic events, resulting in altered timing and progression through meiotic landmarks.

In the context of the *tcx5;6* diploid mutants, we anticipate deviations from the wild-type meiotic timeline due to potential disruptions in chromosomal behaviour and checkpoint control mechanisms. These disruptions are likely to manifest as variations in the duration of meiotic stages, detectable through a comparative temporal analysis with wild-type diploids grown under standard conditions.

The box plot presented in Figure 4.18 illustrates the variation in the duration of meiotic transitions between *tcx5;6* mutants and wild-type controls. We expect any perturbations in chromatin structure and gene regulation due to the *tcx* mutations could result in an extended or abbreviated duration of specific meiotic stages. For example, an extended leptotene or zygotene phase may indicate challenges in chromosomal synapsis, while an abbreviated pachytene stage could suggest premature progression without the completion of necessary recombination events.

Compared to the control, the data from the *tcx5;6* mutants is expected to show increased variability in meiotic duration across different stages. This would reflect the mutants' compromised ability to regulate the progression of meiosis, which is typically tightly controlled in wild-type plants. Furthermore, we may observe a general trend of extended meiotic stages as the mutants grapple with potential deficiencies in chromosomal pairing and segregation.

4.5.5 Conclusion

In conclusion, our evaluation confirms that the Automated Meiosis Progression (aMP) framework represents a significant advancement in plant science, particularly in the automated analysis of live-cell imaging for meiosis progression. Through applying deep learning techniques and innovative tracking methodologies, we have successfully automated the classification and staging of meiocytes in video sequences to stage and form the piecewise timeline of meiosis progression, overcoming the traditional bottleneck of manual image analysis.

The effectiveness of the aMP framework was demonstrated through the analysis of wild-type and tetraploid *Arabidopsis thaliana*, highlighting its capacity to adapt to and reveal the nuances of meiotic progression across different ploidy levels. In tetraploids, the aMP framework detected an extended duration in meiosis-II stages compared to the wild-type, providing new insights into the meiotic behaviour of polyploid plants. Additionally, the framework's extension to

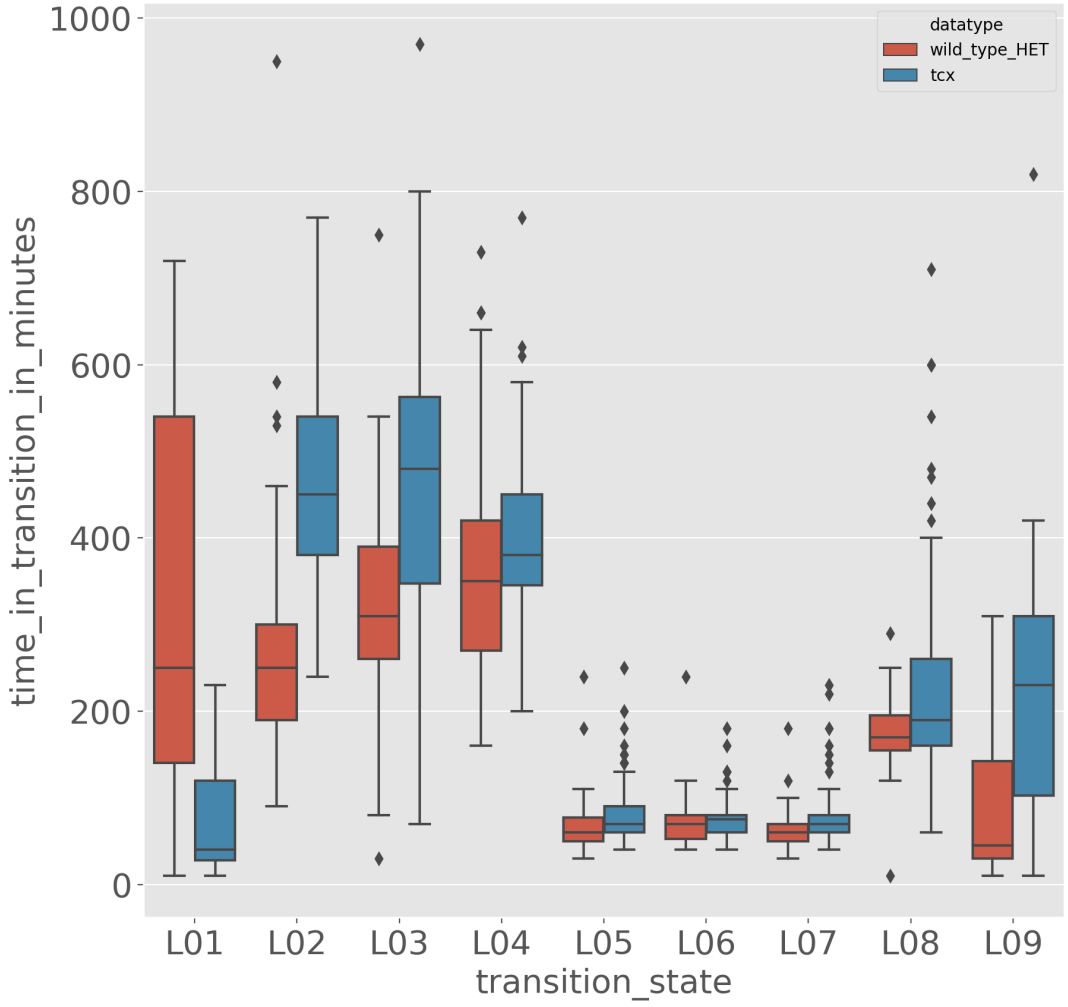

Figure 4.18: ***Comparing tcx mutant to the wild type A.thaliana variants***

analyze *tcx5;6* mutants showcases its versatility and potential for broader applications in plant genetics.

Our research has automated a critical aspect of plant developmental biology and set the stage for further investigations into the genetic and environmental factors that govern meiosis. The aMP framework is a powerful tool for plant science analytics, offering a sophisticated means to study the complexity of meiosis with greater accuracy and efficiency, also creating potential analytical fields for transferring our modular framework, detailed in the next chapter.

Chapter 5

Modularity of amp Pipeline

In this chapter, we delve into the multifaceted applications of the aMP toolkit, a modular system designed for diverse challenges in plant biology. Originally developed for meiosis analysis, this toolkit's versatility extends to tasks such as pollen counting, root length measurement, and foci quantification during DNA double-strand breaks in prophase-I of meiosis. Initially proposed for anther stabilization, the segmentation block demonstrates remarkable adaptability in these applications. The aMP toolkit's modular design is a testament to its plug-and-play capability, allowing for potential enhancements with more specialized feature extractors beyond the EfficientNet family. This is particularly advantageous for complex tasks like BiFC analysis, where accurate chloroplast localization in plant tissue is critical.

Embracing a modular approach aligns with the growing trend of foundational model methodologies, exemplified by systems like Segment Anything from Meta. However, due to hardware limitations, our focus has been on convolutional neural networks rather than next-generation transformers, leading to the development of a unique in-house bioinformatics foundational model. Our approach involves dividing specific tasks, creating variations of the original modules with different models for distinct applications, and ultimately synthesizing a comprehensive model for semantic segmentation across all modules.

We address four key issues in plant biology, starting with each application's objective and specific post-processing requirements.

Firstly, quantifying foci for estimating DNA double-strand breaks (DSBs) during prophase-1 meiosis is pivotal. As detailed in Chapter 2, meiosis halves the nuclear DNA content and facilitates genetic diversity through homologous chromosome recombination. Meiotic recombination starts with DSBs, catalyzed by SPO11, with subsequent processing by the MRE11-RAD50-NBS1 complex. This results in two recombination products: crossovers and non-crossovers. Our study focuses on the number of DSBs in the absence of ZmSDS1 and ZmSDS2 proteins, using *Zea mays* for immunolocalization assays to detect RAD51 foci.

Secondly, the toolkit addresses a crucial aspect of BiFC (Bimolecular Fluorescence Complementation) analysis: the co-localization challenge. The goal here is to detect and quantify protein-protein interactions within living cells accurately. This is achieved through the YFC (Yellow Fluorescent Protein) channel, a primary indicator of these interactions. In BiFC, the interaction of proteins is inferred from the proximity-induced fluorescence of YFP (Yellow Fluorescent Protein) fragments, each attached to a protein of interest. When these proteins interact,

the YFP fragments merge to emit a distinct fluorescence signal. This fluorescence, observed in the YFC channel, is the key to identifying interacting proteins. Additionally, mCherry, a red fluorescent protein, is a transformation control. The presence of mCherry fluorescence confirms the successful transformation of cells and aids in selecting the appropriate cells for BiFC analysis. By focusing on cells exhibiting YFP and mCherry signals, the toolkit enhances the reliability and accuracy of detecting and quantifying protein-protein interactions.

Thirdly, the toolkit tackles the significant task of automated pollen counting, a method pivotal for analyzing pollen viability under various environmental and genetic conditions. The primary objective in this aspect of the toolkit is to enhance the accuracy and efficiency of pollen viability assessment, a process traditionally hampered by manual and labour-intensive methods. The toolkit employs advanced image analysis techniques in automated pollen counting to identify and quantify individual pollen grains from microscopic images. This process addresses the segmentation challenge, where the system distinguishes each pollen grain amidst complex backgrounds, such as other cellular material or the slide surface. The goal is to provide a high-throughput, objective, and consistent approach to counting pollen, essential for studies involving large datasets. This automated process is especially crucial when assessing pollen viability in response to different temperature conditions or following genetic modifications like gene insertion. The toolkit allows for the rapid analysis of many samples, offering significant improvements over manual counting in terms of speed and reduction of human error.

5.1 DNA double-strand breaks (DSBs)

Building upon the modular and versatile capabilities of the aMP toolkit, as introduced earlier, we delve into the specific area of DNA double-strand breaks (DSBs) quantification. This module of the toolkit is developed in response to existing methodologies where foci, particularly RAD51 foci in meiosis, are predominantly quantified using various multipurpose software like Image tool 3.0 [71], softWoRx [72], and ImageJ-FIJI [73, 74], or in some cases, are assumed to have been counted manually [75–78]. Other software mentioned in these studies includes ZEN2012 and Photoshop 7.0. A gap identified in these approaches is the lack of explicit parameters used for quantifying RAD51 foci, leading to the formulation of three important questions:

- What is the size of a focus?
- How bright (intensity) should a focus be compared to the background?
- How are foci distributed over chromosomes, or rather, what is the co-localization overlap percentage of foci on chromosomes?

Our approach, aligned with the toolkit's objective of providing advanced, user-friendly solutions, aims to develop an automated, parameter-free method for RAD51 foci quantification in the maize inbred line A188. We implemented this through a 3-stage process: first, constructing an automated ground truth basis of RAD51 foci in wild-type A188 maize meiocytes using biological prior information; second, benchmarking this process against existing parameterized methods; and finally, employing a convolutional neural network-based segmentation algorithm to eliminate reliance on manually set parameters. This part of the study is crucial, as it contributes to our understanding of the variation in area, intensity, and overlap ratio of foci. It also initiates the development of a standardized definition for RAD51 foci quantification, enhancing the toolkit's applicability in advanced plant biological research.

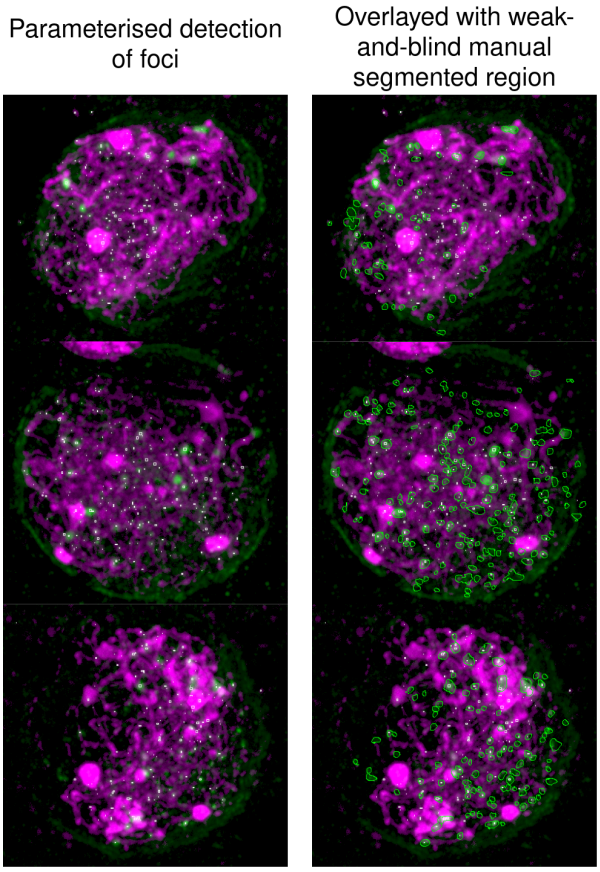

Figure 5.1: Parametric detection of RAD51 foci
Illustration of parametric detection of RAD51 foci through boundary conditions, with weak-and-blind segmentation for process validation. The right column shows manual segmentation contours, and white dots represent automated foci detection. The left image demonstrates clear foci contours alongside automated detections.

5.1.1 Dataset description

In our focus on DNA double-strand breaks (DSBs) quantification, as detailed in the previous section, we utilized a specific dataset to develop and train our quantification model. This model aims to estimate the foci on chromosomes within maize nuclei, a critical task in understanding meiotic processes.

Our dataset comprises 39 confocal microscopic images of wild-type A188 maize meiocyte nuclei. These images are 2D-stacked and feature two channels, each with dimensions of 896×470 pixels and a depth of 16 bits. The images showcase RAD51-stained foci co-localized with DAPI-stained chromosomes, exhibiting varying intensities, essential for accurately assessing foci distribution and density.

Additionally, for inference and comparative analysis, the dataset includes immunolocalization images of both wild-type A188 and the *sds1 sds2-2* mutant meiocytes. These images are crucial for examining the localization of several proteins pivotal for meiotic recombination, such as HEI10, ZYP1, and MLH3. The comparative analysis is structured to focus on two forms:

- Foci, including proteins like RAD51, SDS1-RFP, HEI10, and MLH3.

Parameter Name	Lower Bound	Upper Bound
Intensity of the Focus	Global mean of RAD51 stacks	
Diameter of the Focus	$1\mu m$	
Diameter of the Nucleus	Segmented max project of DAPI channel	$22\mu m$
Occurrence of Foci	2 consecutive stacks	

Table 5.1: Basis for Quantifying Foci in a Nucleus

- Chromosome axis, highlighted by proteins such as ASY1 and ZYP1.

The confocal microscopic images used for network training were acquired using an LSM880 microscope, with samples immersed in oil and magnified 63x. The effective Numerical Aperture (NA) for these images is 1.4, with a stack focus technique applied. The images are defined with pixel sizes of $\Delta z = 381\,\text{nm}$, $\Delta y = 50\,\text{nm}$, and $\Delta x = 50\,\text{nm}$ in each respective direction. The channels are distinctly coloured to aid in identification: the DAPI-stained chromosomes are defined in magenta (`FF00FF`) with an excitation wavelength of 450 nm, while the foci are represented in green (`00FF00`), having an excitation wavelength of 548 nm.

5.1.2 Parametric Detection of Foci on Chromosome

In this subsection, we elaborate on the parametric model developed to automate the detection of foci, specifically RAD51-stained double-strand breaks, co-localized with chromosomes in maize meiocyte nuclei. Our methodology, linked to the broader framework of the toolkit, focuses on defining precise boundary conditions based on pre-formulated queries and the inherent characteristics of the maize meiocyte nucleus.

Considering the average diameter of the maize meiocyte nucleus at $\approx 20\mu m$ [79], we set a constrained region of interest for locating foci, given their emergence from meiotic recombination. The lack of studies addressing foci diameter led us to adopt the human chromosomal cross-sectional diameter during prophase of $1\,\mu m$ [80] as our lower boundary condition. Furthermore, the intensity of foci, generally higher than other image elements, necessitated setting a lower intensity threshold based on the global mean of the combined z-stacks of the cellular image. The RAD51 foci identified are then correlated with DAPI-stained chromosomes to determine the number of foci and clusters, thereby quantifying double-strand break formation. These clusters form the basis for human intervention justification on the images' maximum projection, enabling weak segmentation of foci for serializing the machine learning dataset.

We employed a channel-agnostic methodology in the automated process for generating data on double-strand breaks. This involved processing RAD51-stained foci and DAPI-stained chromosome (chx) stacks through an image conversion block. During preprocessing the training dataset, we applied a maximum projection technique to the region of interest (ROI) in the 2D-stacked images. We transformed the original 16-bit images, with dimensions of 489×407 pixels and two channels, into 8-bit RGB images.

The chromosomes, highlighted by DAPI staining, were then processed through a Gaussian-adaptive thresholding technique applied to each stack individually. This approach was optimized with a window size of 21 pixels to delineate the nuclear boundary accurately. Following this, a median filter with a window size of 11 pixels was employed to remove background

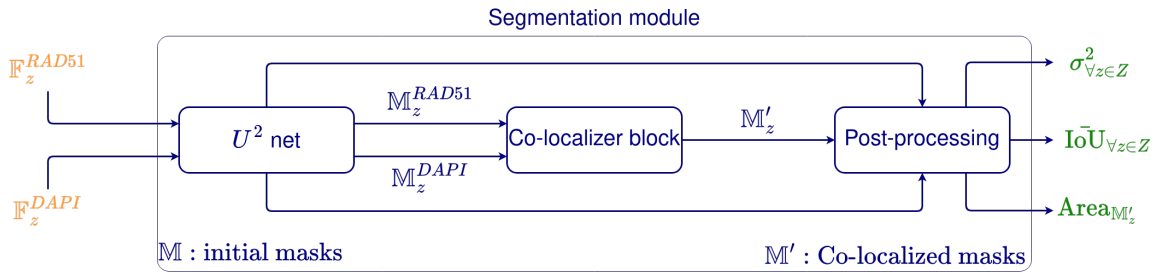

Figure 5.2: Foci detection process flow
The aMP segmentation module used in segmenting DAPI-stained chromosome and RAD51-stained foci illustrates the inference flow for quantifying double-strand breaks based on three major criteria: overlap percentage, size, and intensity variation.

speckles and noise. This filtration step was crucial in defining the nuclear boundary, effectively differentiating the background from the foreground. To facilitate the subtraction of the background, we adopted a principle based on 'one standard deviation below the mean', denoted as $\mathbb{Z} = -1$, which can be mathematically represented as:

$$\mathbb{Z} = \frac{\text{datapoint} - \text{mean}}{\text{standard deviation}} \tag{5.1}$$

After applying maximum projection and performing background subtraction on a stack-wise basis, the next step in our process involves analyzing the chromosome stack. This stack is processed to determine the global threshold required for extracting the region of interest (ROI) from the RAD51-stained image stacks. To efficiently identify the local maxima, which signify potential foci locations, we employed a kernel size of 20×20 pixels. This size was selected to align with the lower bound estimate of chromosome diameter, approximately $1\mu m$. Additionally, a hop size of 20 pixels was utilized to navigate through the image, a strategic choice to minimize computational time.

Once the potential foci points were identified, they were subjected to clustering analysis using the Voor-Hees algorithm. This algorithm was chosen for its effectiveness in grouping data points based on the Euclidean metric, a decision that further contributed to reducing the overall timing complexity of the process. The clusters formed through this algorithm are considered as the estimated foci. This estimation is crucial as it quantifies double-strand breaks in the nuclei. Using this methodical approach, we efficiently and accurately identified and quantified the foci, thereby providing valuable insights into the occurrence and characteristics of double-strand breaks.

5.1.3 Preparation of the Ground Truth

In the process of developing a ground truth dataset for training our machine learning system in foci quantification, we implemented a two-step procedure on the maximum projections of A188 maize nuclei images:

- The first step involved segmenting the most evident foci. This segmentation was based directly on the visual characteristics of the foci, due to the absence of a pre-established

statistical definition for foci dimensions and intensities.

- The second step entailed overlaying these manually segmented foci with those our automated detection system identified. This step was critical to verify the effectiveness of our automated method.

The ground truth dataset was prepared with input from immuno-experts who identified foci regions without using automated positional indicators. Their segmentation focused on regions where RAD51 staining was noticeably different from the background. This manual segmentation process was essential for ensuring the accuracy of the automated system.

We then compared the manual segmentation results with the outcomes of the automated foci quantification process. This comparison was essential to validate the automated system's accuracy and refine the training dataset by eliminating images with poor quality, such as those with smeared backgrounds or inadequate immuno-staining in the cytoplasm.

Moreover, this detailed approach allowed us to gather data on foci size and number variations under different conditions, including in mutant samples and with varying immuno-staining methods. This information is crucial for the reliable training of a machine-learning model that can accurately quantify foci in maize meiocytes.

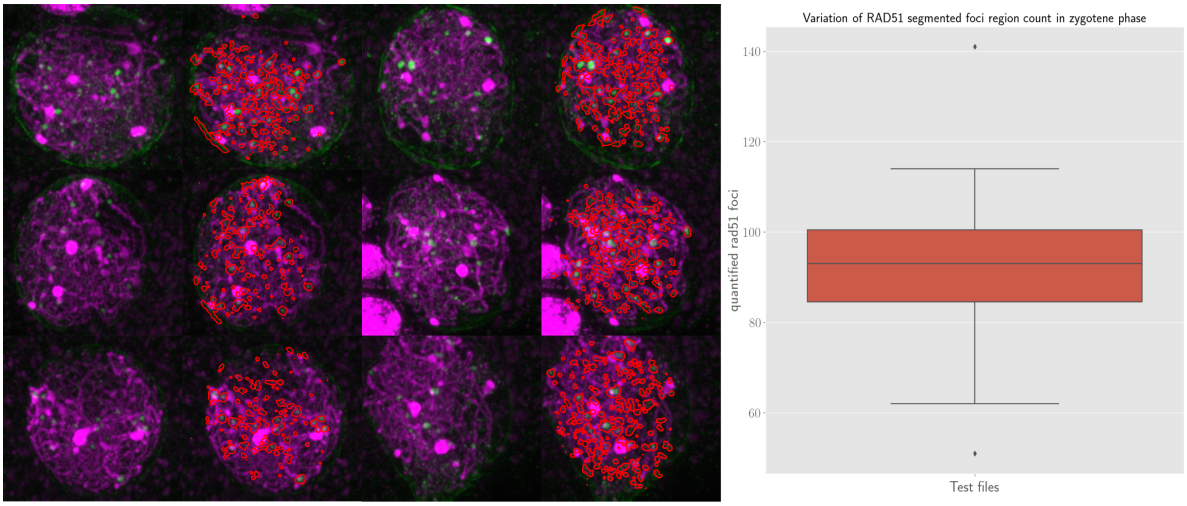

Figure 5.3: Localized foci visualization on maximum projected images.

	Maximum	Average	Minimum
Foci Count	359	209	133
Overlap w.r.t. Clusters	43%	18%	3%
Overlap w.r.t. Mask	50%	30%	7%

Table 5.2: Comparison of Parametric Foci Generation with Manual Segmentation

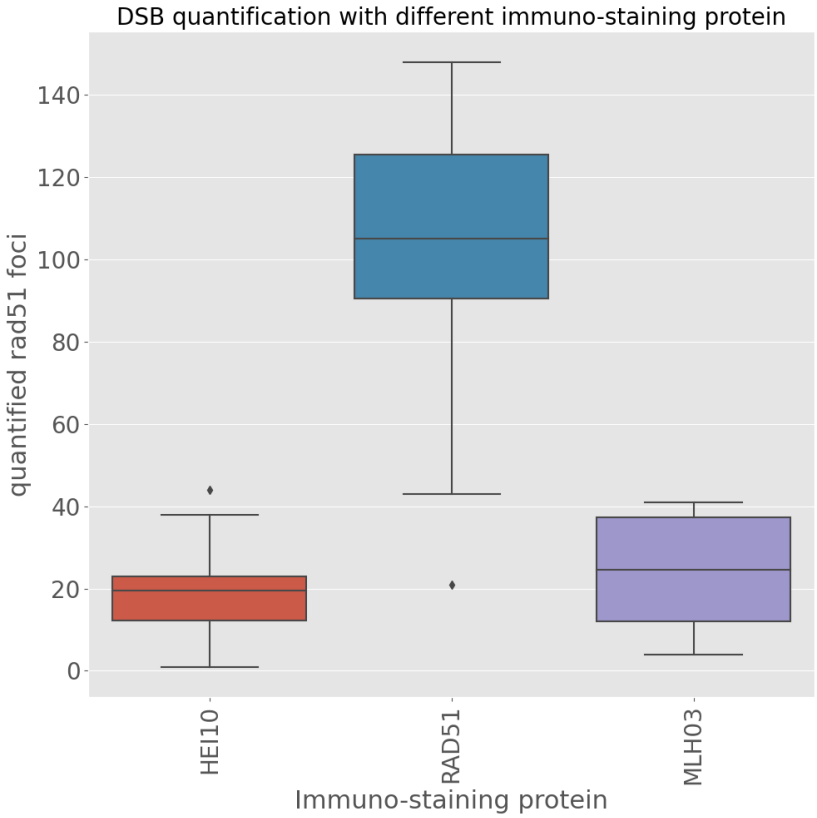

Figure 5.4: Quantifying localized foci across different immunostainings on maximum projected images.

5.1.4 Enhancements in Foci and Chromosome Segmentation

The configuration of the segmentation network aligns with the aMP segmentation module, depicted in Fig.5.2. It involves modifying the post-processing steps to include assessments of Intersection over Union (IoU), area, and intensity variation, enhancing the accuracy of the segmentation.

To precisely evaluate the overlap of foci with chromosomes, we implemented distinct training processes for chromosome and foci segmentation. This methodological distinction was crucial for accurate overlap assessment and effectively filtering out images unsuitable for reliable double-strand break quantification analysis.

Building upon the ground truth data established in the previous subsection, we enhanced our segmentation approach to refine the analysis of RAD51-stained double-strand breaks in maize nucleus images. This advancement involved two major steps: firstly, the integration of a deep learning-based segmentation network, and secondly, the implementation of a manual verification process to validate the automated findings.

Our parametric method for counting RAD51 foci yielded an average of approximately 209 foci per maize nucleus, aligning with previous findings in the zygotene phase. However, the lack of clarity in previous studies regarding the exact methods of foci counting (either based on maximum projection or individual stack analysis) limited our ability to draw firm conclusions. To address this, we manually segmented foci without considering chromosome position and

compared these with automated detections, achieving a maximum overlap of 50% and an average of around 30%.

To further validate our approach, we applied a multiplicative-attention-based, deep neural network for segmentation, focusing on intensity differences and respecting nuclear boundaries. This network was trained on maximum projection images and tested on individual stacks, as illustrated in Fig.,5.3 and Fig.,5.4.

Our tests on different A188 maize nucleus images, detailed in the bottom row of Fig.,5.6, revealed variations in segmented region size, mean intensity, and area-based outliers. These findings highlight the need for a comprehensive foci detection threshold, which can assist in filtering out non-contributory image slices.

The deep-learning-based approach proved more effective in standardizing foci quantification than traditional parametric methods. We could quantify double-strand break formation with higher accuracy by employing a boundary condition based on the chromosomal diameter. This methodology, tested on a diverse range of immunostained samples, demonstrates its robustness and versatility, making it a valuable tool in analysing double-strand breaks in plant biology.

5.2 BiFC Analysis

Following our exploration of DNA double-strand break (DSB) segmentation and co-localization processes, we expanded our focus to BiFC (Bimolecular Fluorescence Complementation) analysis. This extension leverages the insights gained from the previous section, particularly in signal ratio analysis between YFP (Yellow Fluorescent Protein) and mCherry signals in Tobacco leaf. BiFC is the technique to evaluate molecular interactions. In the BiFC experiment, we use the non-fluorescent N-terminal and C-terminal fragments of a fluorescent protein. Each

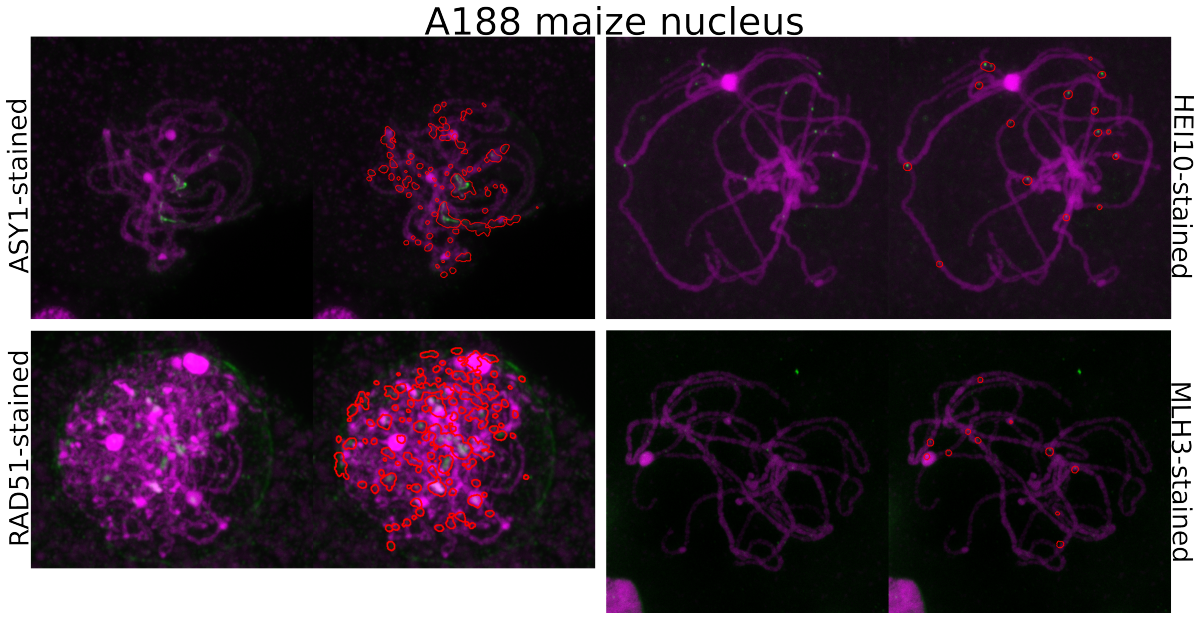

Figure 5.5: Visualization of localized double-strand breaks for various immunostainings.

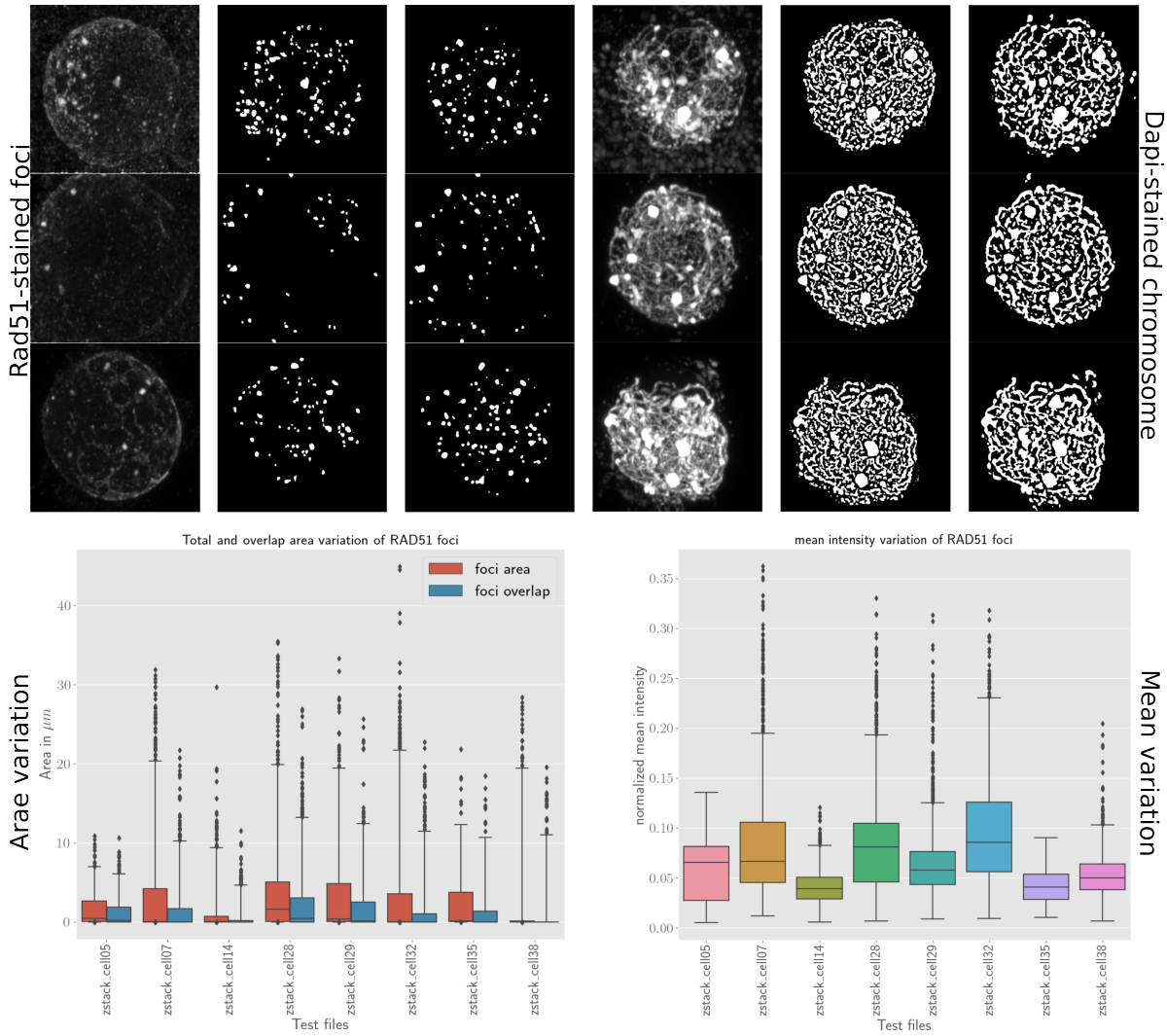

Figure 5.6: Network validation for foci quantification
The top row shows examples of test images with reference labels and the network prediction. We can observe the benefit of Unet-2D-based network estimation as the network learns about the intensity variation to the surroundings. The bottom row shows the variation of foci and its overlap with chromosomes and the intensity variation.

of the fragmented non-fluorescent proteins is tagged with proteins of interest. These fragmented fluorescent proteins complement and generate fluorescence if these two proteins interact. We can validate the interaction based on the fluorescent intensity.

We identified potentially interacting molecules with our target protein, JASON. JASON is the important protein for proper chromosome segregation during meiosis II [81, 82]. It has already been reported that JASON has a function of generating organelle bands, which plays a role in preventing the fusing of two spindles in meiosis II [83]. Since the potential interacting proteins with JASON are membrane proteins localized on the chloroplast, we must evaluate the BiFC signal on the chloroplast membrane. The primary objective of our BiFC analysis was to segment chloroplasts, marked with mCherry and auto-fluorescence, and to segment the YFP channel. Crucially, chloroplast regions of interest (ROIs) not overlapping with YFP signals were discarded. This approach aligns with the earlier method of segmenting foci, where the objective was to co-localize specific signals within the cellular structure.

Our methodology was as follows:

- **Cell Selection**: We first identified cells exhibiting the mCherry signal, indicating successful transient transformation, qualifying them for subsequent BiFC (YFP) signal evaluation.
- **Chloroplast Segmentation**: Chloroplasts were identified predominantly by their autofluorescence. Only cells with a clear mCherry signal underwent chloroplast segmentation to ensure transformation efficiency.
- **BiFC (YFP) and mCherry Ratio Measurement**: We then measured the YFP intensity specifically on the chloroplast membranes. Concurrently, we quantified the mCherry signal to establish the YFP/mCherry ratio, providing an index of protein-protein interaction strength.

We adopted a channel-wise approach for ground truth generation, similar to the FOCI mask creation. YFP channel ground truths were generated analogous to chromosome mask generation. We introduced adjustments in image preprocessing, such as contrast enhancement and dilation, which significantly improved segmentation detection, with a dice score reaching approximately 0.73. The application of dilation was identified as a critical hyperparameter, particularly for smaller ROI in segmentation challenges. The dilation window was varied based on ROI statistics in the ground truth data.

Our dataset for BiFC analysis comprised 76 images, each with 2 channels of dimensions 1024×1024 pixels and a depth of 8-bit gray-scale. This dataset included both positive and negative control samples for various protein-protein interactions. The inference dataset (see Table. 5.3) encompassed various samples representing different protein-protein interactions relative to the control samples.

5.2.1 BiFC segmentation results and post-processing analysis

The segmentation module (Fig. 5.7), similar to that used in foci quantification, integrated a post-processing block tailored for BiFC analysis. Network predictions, illustrated in (Fig. 5.8),

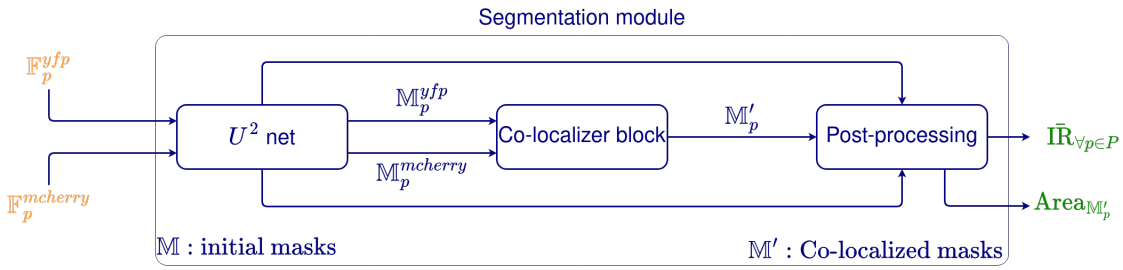

Figure 5.7: *aMP segmentation module for BiFC segmentation.*
The process begins with input fluorescence channels (YFP and mCherry) undergoing segmentation via a U2-Net-based architecture. Initial segmentation masks for each channel are then co-localized to identify overlapping regions. Post-processing refines these co-localized masks to output quantifiable metrics such as the intensity ratio (IR) and area of the YFP signal, critical for analyzing protein-protein interactions in BiFC assays.

Sample ID	Protein 1	Protein 2
S01	JAS-cYFP	nYFP-TOC132
S03	JAS-cYFP	nYFP-TOC120
S05	JAS-cYFP	nYFP-TOC33
S06	JAS-cYFP	nYFP-TOC34
S10	JAS-cYFP	JAS-nYFP
S12	JAS-cYFP	PS1-nYFP
S13	cYFP-TOC33	PS1-nYFP
S16	cYFP-TOC33	nYFP-PS1
S14	cYFP-TOC132	PS1-nYFP
S17	cYFP-TOC132	nYFP-PS1
S15	JAS-cYFP	nYFP-PS1
S18	cYFP-JAS	nYFP-JAS
S19	JAS-cYFP	nYFP-JAS
S20	JAS1-130aa-cYFP	nYFP-TOC33
S21	JAS131-250aa-cYFP	nYFP-TOC33
S22	JAS251-365aa-cYFP	nYFP-TOC33
S23	JAS365-480aa-cYFP	nYFP-TOC33
S25	cYFP-JAS131-250aa	nYFP-JAS
S26	cYFP-JAS251-365aa	nYFP-JAS
S27	cYFP-JAS366-480aa	nYFP-JAS

Table 5.3: BiFC inference dataset
Detailed tested protein-protein interactions

revealed an enhanced detection of the chloroplast region due to the network's multiplicative attention mechanism. This enhancement suggests an advanced feature extraction capacity that could benefit from an expanded dataset in future studies.

Post-processing techniques like dilation and contrast enhancement were pivotal in accurately localizing smaller regions of interest and mitigating noise artefacts from image preprocessing, resulting in clearer segmentation outputs. Our primary objective was to compute the ratio of YFP to mCherry signal intensities within the co-localized chloroplast mask, facilitating the quantification of protein-protein interactions. The intensity ratio (IR) is calculated using the following equation:

$$\bar{IR} = \frac{I(\text{YFP channel})}{I(\text{mCherry channel})} \tag{5.2}$$

This ratio quantitatively measures interaction strength within the BiFC assay framework. Analytical results depicted in Fig. 5.9 demonstrated that in the negative control samples, where no interaction between JASON and the cytochrome b5 family of proteins (localized to the outer membrane of the chloroplast) was expected, no YFP signal was detected despite the presence of the mCherry signal. Conversely, positive controls involving known interacting proteins TOC34 and TOC132, as well as TOC34 and TOC120, exhibited strong YFP signals both in the cytosol and on the chloroplast membrane, with a significantly higher YFP/mCherry ratio than the negative control, validating the effectiveness of our BiFC analysis and experimental setup.

Subsequent experiments with our samples revealed a discernible interaction between TOC33

and JASON, indicated by a higher YFP/mCherry ratio. However, no interaction was detected with other combinations involving JASON. Further testing for dimerization suggested that N-terminally tagged JASON molecules could interact with each other, hinting at JASON's role in facilitating organelle clustering during meiosis.

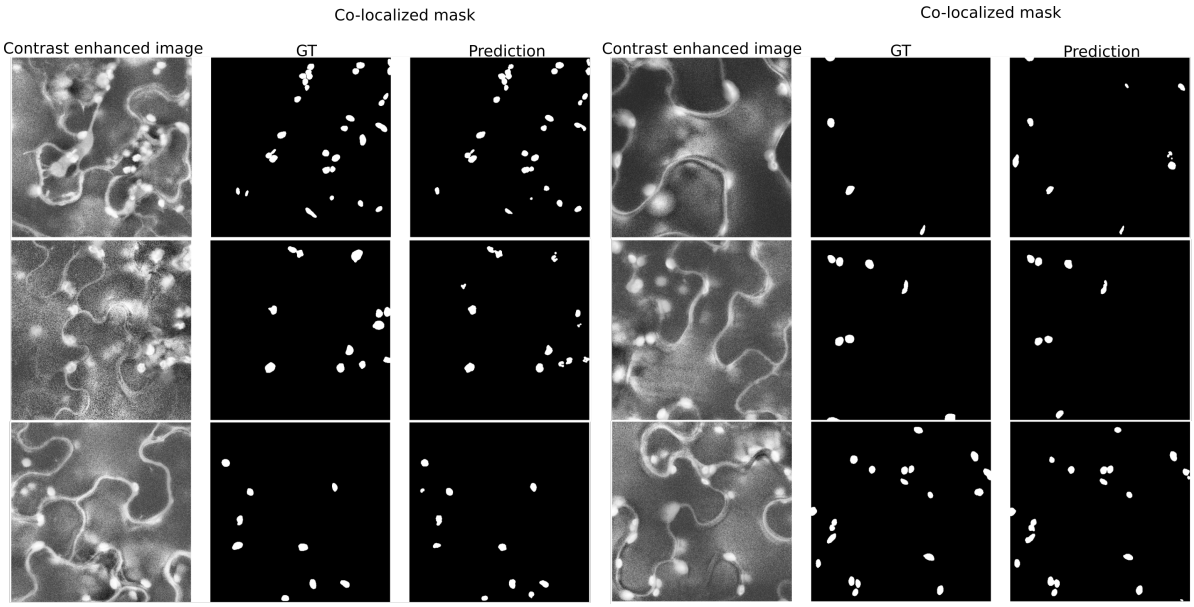

Figure 5.8: ***Comparative visualization of segmentation in BiFC analysis.*** The left column displays contrast-enhanced dilated original microscopy images of tissues undergoing protein-protein interactions. The middle column shows the corresponding ground truth mask highlighting co-localised chloroplast, and the right column presents the post-processed segmentation prediction, where co-localized regions of interest have been isolated for quantitative analysis.

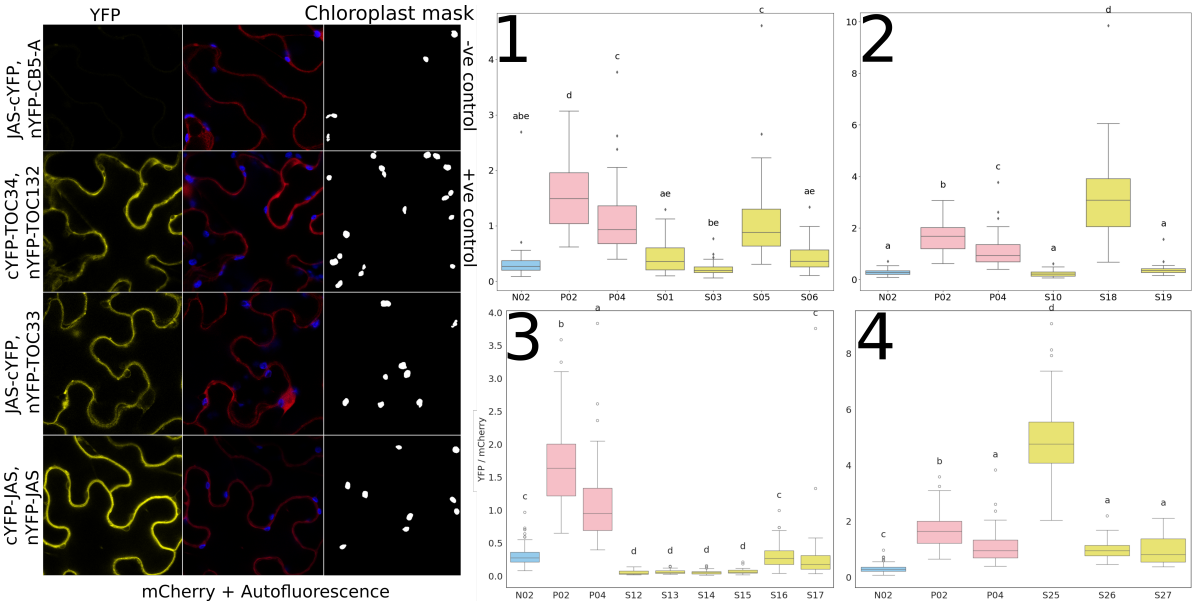

Figure 5.9: ***BiFC inference.***
1. Jason interacts with TOC33 in BiFC assays; 2. Jason can dimerize, 3. PS1 doesn't interact with JASON nor TOCs in BiFC assays, 4. The part of its 131-250aa mainly derives JAS dimerization

5.3 Pollen counting

The segmentation module of our toolkit has been further refined to address the task of pollen counting. This process focuses predominantly on the post-processing stage, where the primary goal is to accurately isolate and enumerate pollen grains while effectively distinguishing them from debris and other non-target elements commonly present in scan images. The enhancements in the pollen segmentation module mirror the strategies employed in anther segmentation, focusing on identifying and segmenting pollen grains. The key contributions to this enhanced segmentation process are twofold:

1. The introduction of random noise and extraneous elements into the training images, simulating various background scenarios. This augmentation enhances the model's robustness, enabling it to isolate pollen grains reliably across diverse and unpredictable background conditions that represent real-world scenarios. Mathematically, this can be represented by augmenting the original image $\mathbb{F}_{\text{pollen}}$ with a noise function η, such that the augmented image $\mathbb{F}_{\text{pollen aug}}$ is given by:

$$\mathbb{F}_{\text{pollen aug}} = \mathbb{F}_{\text{pollen}} + \eta(\alpha, \beta) \tag{5.3}$$

   where η is a function introducing noise or random elements with parameters α and β controlling the type and intensity of the augmentation.

2. The watershed algorithm [84], is used for the accurate counting and separation of overlapping pollen grains. This algorithm conceptualizes the pixel values of a generated mask as a topographical landscape, wherein the 'flooding' of basins is initiated from local minima within the mask image, also known as markers. As the process progresses, the basins expand until they converge on the watershed lines - boundaries between regions associated with different markers. The mathematical formulation of the watershed transformation is given by:

$$\mathbb{M}'_{\text{watershed}} = \text{Watershed}(\mathbb{M}, \mathbb{M}_{\text{markers}}), \tag{5.4}$$

   where $\mathbb{M}$ denotes the initial segmentation mask, and $\mathbb{M}_{\text{markers}}$ represents the set of markers that guide the segmentation process by defining the starting points of the 'flooding' basins.

   Post application of the watershed algorithm, the radius r of each segmented pollen grain is derived from its area A. Assuming the grains are approximately circular, the radius can be calculated using the area of a circle $A = \pi r^2$, hence the radius is determined by the equation:

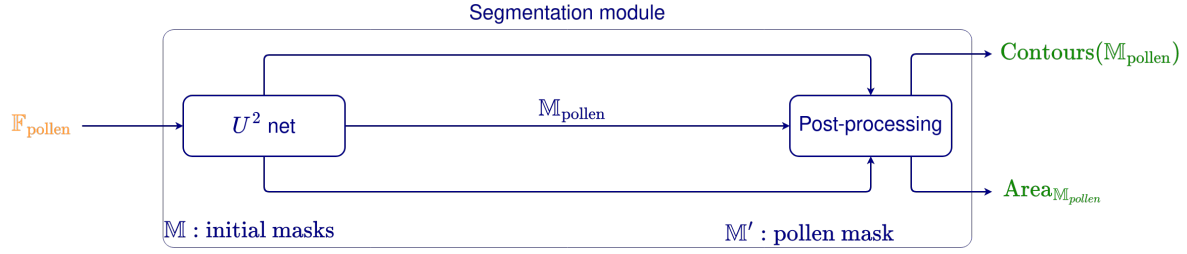

Figure 5.10: ***aMP segmentation module for pollen count.***

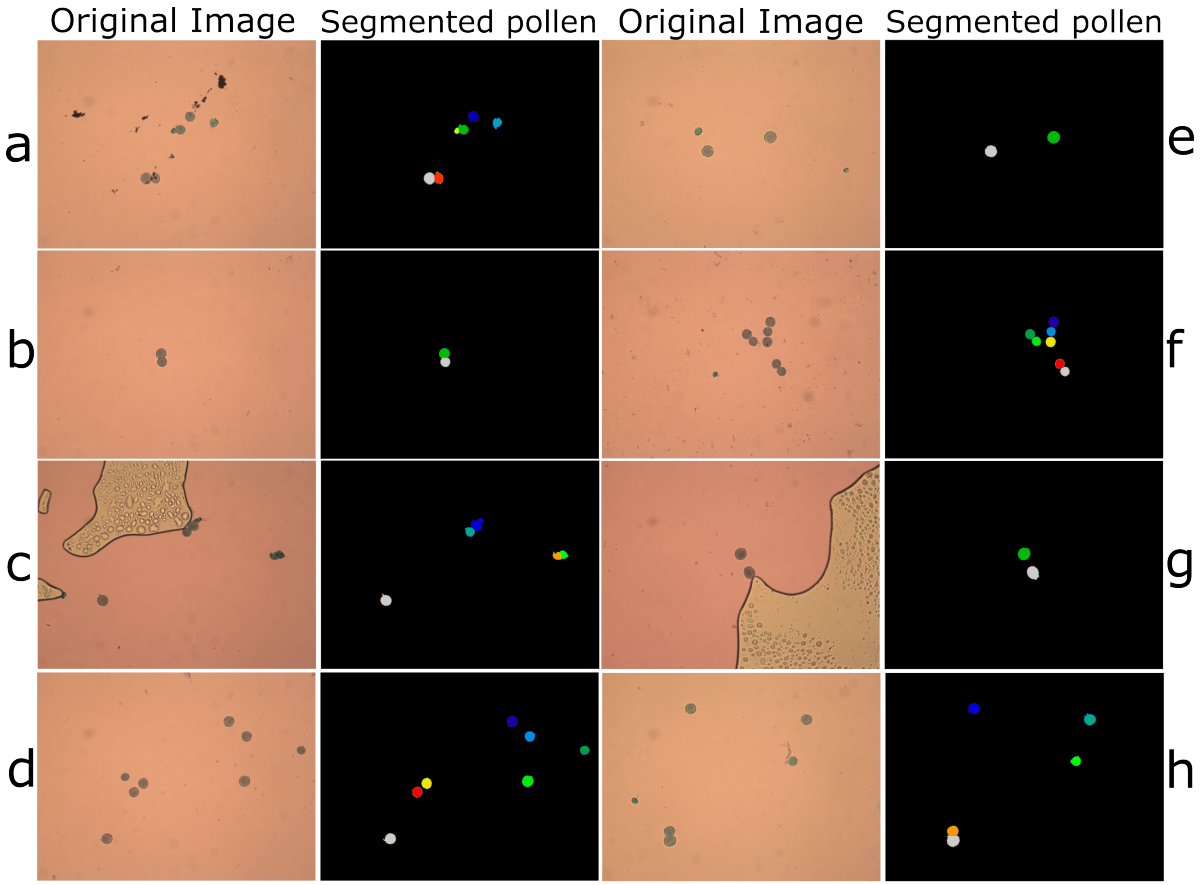

Figure 5.11: ***Pollen segmentation with post-processed results.***

$$r = \sqrt{\frac{A}{\pi}}, \tag{5.5}$$

, where A is the area of a pollen grain extracted from the segmented mask $\mathbb{M}'_{\text{watershed}}$. This radius computation facilitates the analysis of the size distribution of pollen within a given sample, providing valuable insights for botanical research.

Our pollen dataset comprises 94 images of *Arabidopsis thaliana* (Col-0) sourced from four distinct plants, which include samples from three flowers each and three separate branches. Each image in the dataset has a resolution of 2584×1936 pixels. The inference dataset is curated from a range of heat shock experiments that span from five to nine days, providing a diverse set of conditions for analysis.

Ground truth annotations were meticulously generated following the methodology established for FOCI mask creation, with manual intervention employed to remove extraneous material such as dirt and other non-pollen tissue elements. During preprocessing, we have specifically introduced a random background to the images, a step that significantly contributes to the robustness of the network. This preprocessing simulates variable staining conditions and potential foreign contamination, commonplace in real-world scenarios. The effectiveness of this augmentation is quantitatively demonstrated by the network's performance, achieving a dice score of approximately 95%. This high level of accuracy underscores the network's capability to segment pollen effectively, validating its potential utility in experimental contexts characterized

by staining variability and contaminant presence.

5.3.1 Segmented pollen and quantification

Fig. 5.11 shows a set of microscopy images and their corresponding segmentation analyses are presented, showcasing the effectiveness of watershed post-processing in distinguishing and enumerating individual pollen grains. The first and third columns display the original microscopy images of the pollen samples. These images capture the pollen grains amidst various background elements, including potential debris and artefacts inherent to the sample or the imaging process. The second and fourth column provides a clearer visualization of the segmentation against the original image backdrop, confirming the accuracy of the segmentation and the appropriateness of the identified regions for further morphological analysis. Each segmented and colour-highlighted pollen grain is accompanied by a numerical identifier, which corresponds to measurements of the grain's radius, denoted in pixels or micrometres, depending on the calibration of the imaging system.

The successful segmentation and radius estimation (Fig.5.12) of each pollen grain underscore the robustness of the watershed post-processing technique in resolving complex overlapping structures. These results are instrumental for further quantitative analyses, such as assessing pollen viability and morphological characterization of pollen.

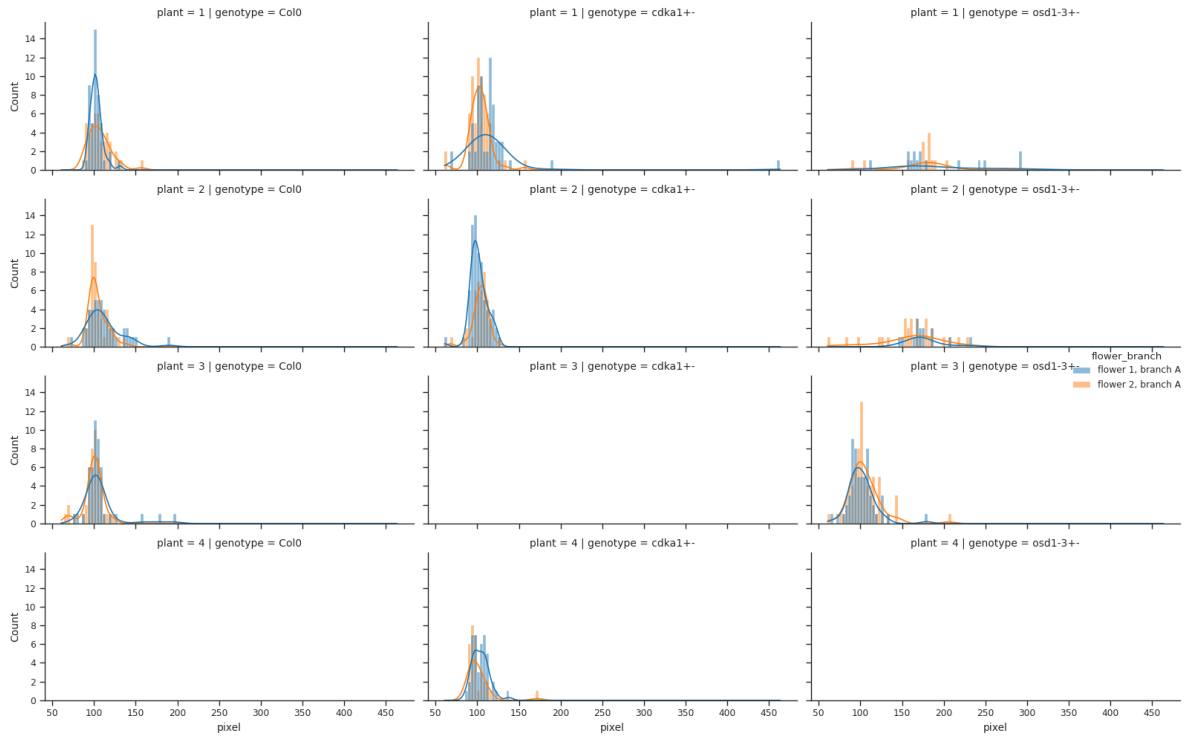

Figure 5.12: ***Control sample at*** $21°C$
Control sample histograms representing the baseline pollen count distribution for standard 'Col0' genotype pollen grains, critical for comparative pollen viability studies with different duration heat-shock treatment.

5.4 Conclusion

The aMP framework is designed to handle a variety of segmentation tasks with remarkable ease adeptly. This framework has rapidly established itself as a pivotal baseline, paving the way towards developing a foundational model tailored for plant physiological analysis. Its primary strength lies in its efficiency, particularly when augmented with application-specific post-processing plugins, which significantly enhance its versatility and applicability across different scenarios.

Looking forward, the aMP framework exhibits tremendous potential for integration with large language models (LLMs). Such an amalgamation could revolutionize the computational analysis landscape by streamlining processes and expanding the breadth of application variations it can accommodate. The envisioned synergy between the aMP framework and LLMs opens up exciting avenues for research and application, promising to elevate plant physiological analytics to unprecedented levels of precision and efficiency. As we continue to refine and expand upon this framework, we anticipate contributing a robust tool that will significantly advance the field, offering researchers and practitioners a sophisticated yet user-friendly analytical instrument capable of tackling the complex challenges inherent in plant physiology.

Chapter 6

Discussion and Future Work

In exploring meiotic cell division dynamics within *Arabidopsis thaliana*, our study has established a foundation for an advanced automated imaging system. This system not only addresses the analytical challenges of today but also paves the way for future innovations in the field by integrating automated microscopy, cell signalling analysis, and detailed examination of key cellular structures has created a robust framework for further investigation.

Our automated microscopy module significantly advances toward real-time, precise observation of meiotic processes. This technology streamlines the imaging process by targeting specific meiotic stages and introduces the possibility of developing real-time analytical tools. The capability to operate in real-time opens avenues for implementing generative AI applications that can adapt and respond to the dynamic changes observed during meiosis, enhancing the system's utility and efficiency.

Understanding the synchrony in division progression among meiocytes is crucial for unravelling the mechanisms of meiosis. The system's ability to capture and analyze division events in detail provides a solid foundation for deeper exploration into the signalling pathways that regulate this process. Future efforts will aim to dissect the complex cell signalling networks, ensuring the coordinated progression of meiocytes through meiotic stages, offering insights into the regulation of meiosis and potential points of intervention for research and therapeutic applications.

The precision with which our system identifies and analyzes structures-from nuclear envelope breakdown (NEB) to anaphase onset and the alignment of kinetochores with microtubules, demonstrates its capability to uncover subtle yet critical information. These insights are instrumental in understanding the physical and molecular dynamics during chromosome segregation. Future assessments will leverage the system's high-resolution imaging and analysis capabilities to delve deeper into these phenomena, uncovering new dimensions of meiotic cell division.

Moreover, the adaptability of our framework is evidenced by its current modification for analyzing maize meiosis progression. This extension demonstrates the system's versatility and highlights its potential applicability across a wide range of species and genetic contexts.

A promising avenue for future exploration is the development of deep learning (DL)–generated 3D structures from 2D stacked images using Generative Adversarial Networks (GANs) and interpolation techniques. This approach aims to revolutionize our understanding of cellular

architecture and dynamics by providing a comprehensive, three-dimensional view of meiotic cells, thereby overcoming the limitations of traditional two-dimensional imaging.

In conclusion, our study introduces an automated system for detailed analysis of meiotic progression and lays the groundwork for developing generative AI applications in this domain. The continued refinement and expansion of this system, its application to diverse species and the exploration of advanced DL techniques significantly advance our understanding of meiosis progression in different phases.

Appendix A

Supplementary results

A.1 Pollen diameter for heat shock experiments

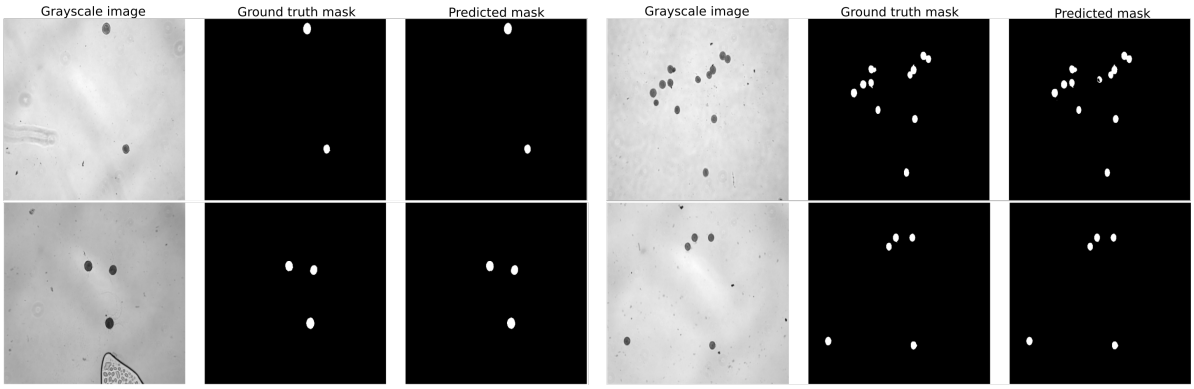

Figure A.1: ***Network evaluation on test images.***

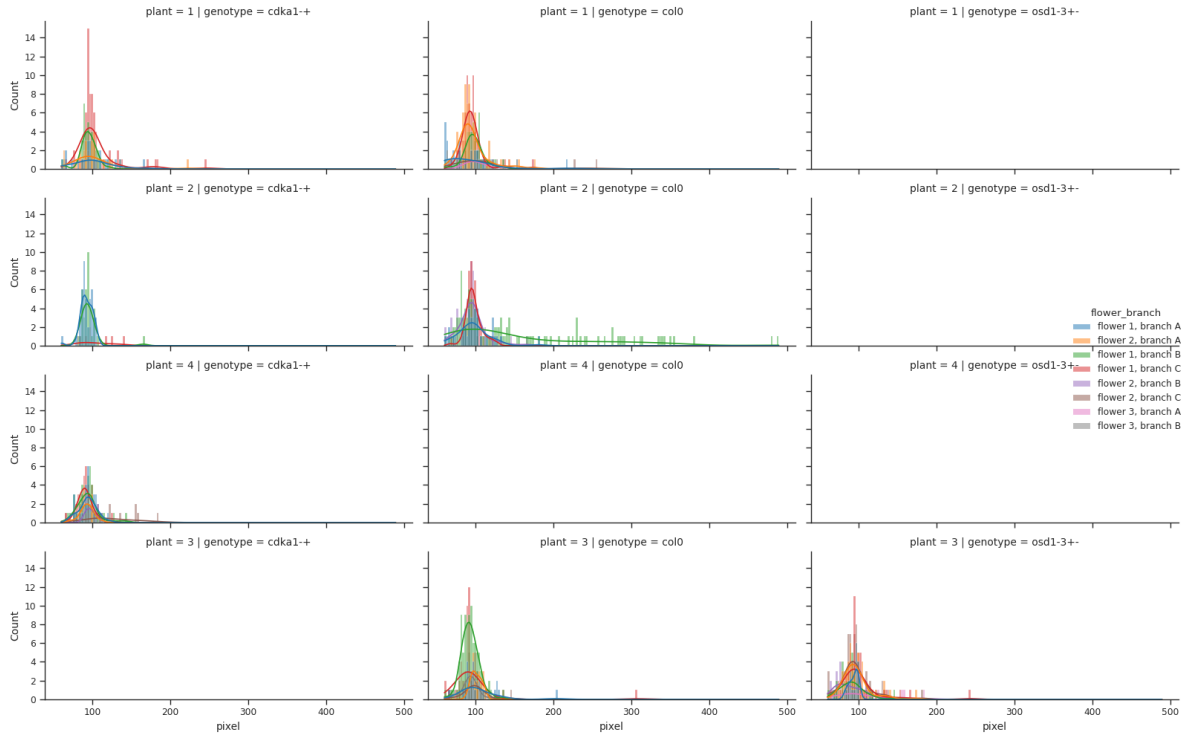

Figure A.2: 6 ***days heat exposure (DAHS)***
Pollen count viability histograms for 'cdka1++' genotype pollen grains at 6 days heat exposure (DAHS).

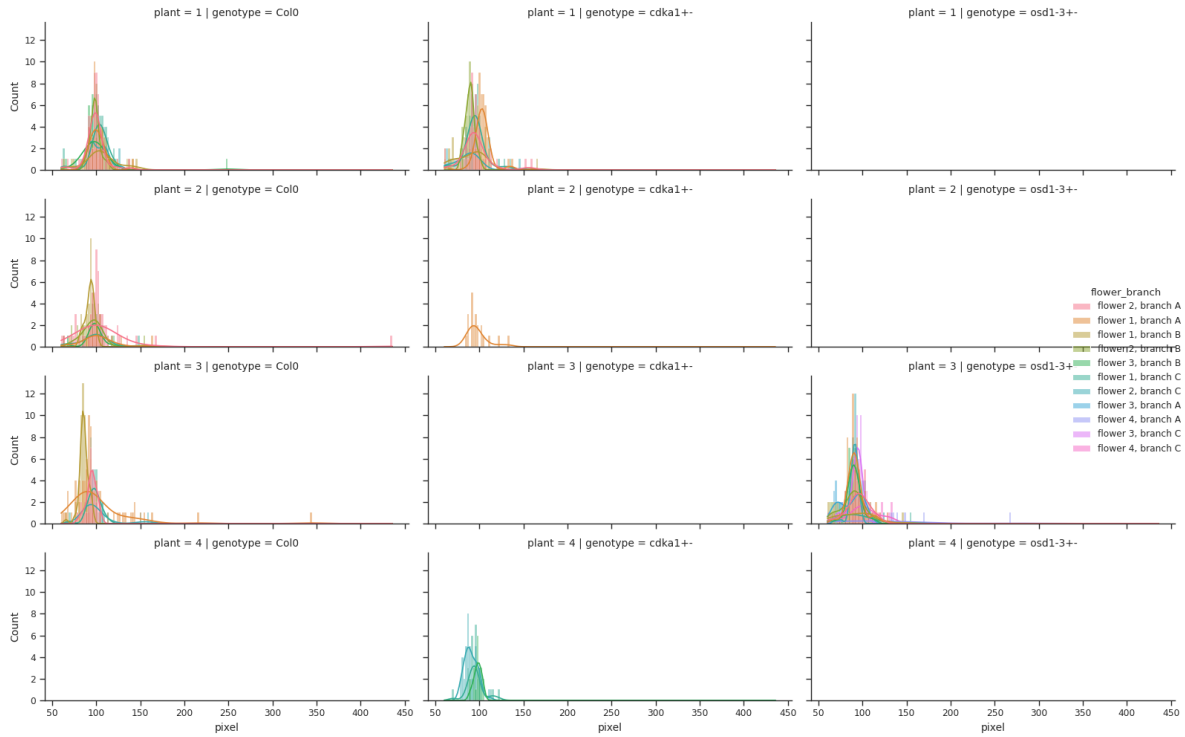

Figure A.3: 7 ***days heat exposure (DAHS)***
Pollen count viability histograms for 'osd1-3++' genotype pollen grains at 7 days heat exposure (DAHS).

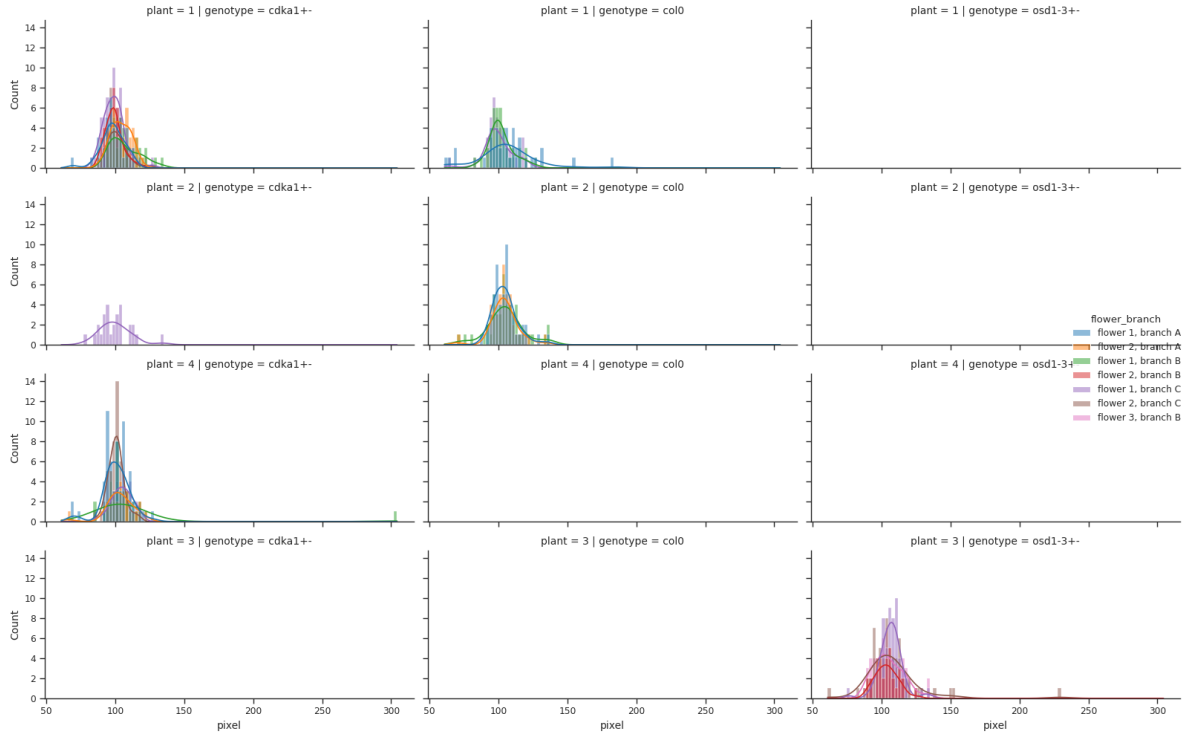

Figure A.4: 8 ***days heat exposure (DAHS)***
Histograms contrasting the pollen viability intensity distributions among various genotypes at 8 days heat exposure (DAHS).

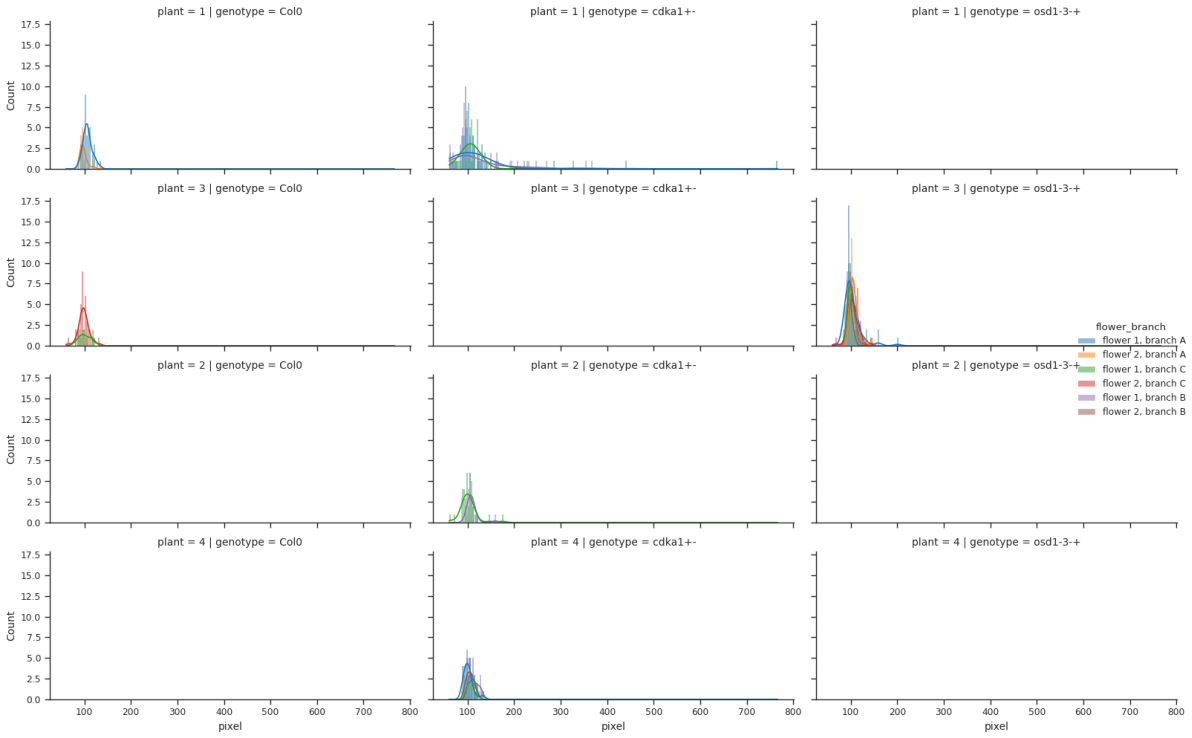

Figure A.5: 9 ***days heat exposure (DAHS)***
Boxplot distribution of pollen grain sizes measured in microns for multiple genotypes at 9 days heat exposure (DAHS), emphasizing the genotype-specific size variations observable in mutant versus 'Col0' strains.

List of Abbreviations

A	
ATM	Ataxia-telangiectasia mutated
AI	Artificial Intelligence
B	
BiFC	Bimolecular Fluorescence Complementation
C	
CS	Cell shape
CNN	Convolutional Neural Network
D	
DNA	Deoxyribonucleic acid
DAHS	Days After Heat Shock
DL	Deep Learning
E	
EB	Entwicklung Biologie
F	
FP	False Positives
FN	False Negatives
FC	Fully Connected
G	
GFP	Green Fluorescence Protein
GT	Ground Truth
H	
HTE	Hasibe Tuncay Elbasi
HT	Hasibe Tuncay
HET	Hasibe Elbasi Tuncay
I	
IoU	Intersection over Union
IDSW	Identity Switches
IPM	Institut für Pflanzenwissenschaften und Mikrobiologie

J

JDJB	Joke de Jaeger Braet

K

KINGBIRG	Kleisin IN Green microtuBules In ReD

M

mAP	Mean average precision
MT	Microtubular array
MOTA	Multiple Object Tracking Accuracy
MOTP	Multiple Object Tracking Precision
MP	Maria Prusicki
MvdH	Max van der Heide

N

NP	Nucleus position
NoP	Nucleolus position
NCC	Normalized Cross-Correlation

R

RC	Chromatin expressed with REC8 marker
REC8	Meiotic Recombination Protein 8
RFP	Red Fluorescence Protein
RSS	Restricted Search Space

T

T	Tapetum
TP	tetraploid
TCX	Tesmin/TSO1-like CXC

U

UHH	Universität Hamburg

W

WT	Wild-type

Y

YH	Yuki Hamamura

List of Software and Platform

Name	Version	URL	Comment
Tensorflow	2.13	https://www.tensorflow.org/	Machine learning library
Python	3.11	https://www.python.org/	Supported version
HuggingFace	-	https://huggingface.co/spaces/arcan3	Deployed system
Git	-	https://gitlab.rrz.uni-hamburg.de	Version control

Table A.2: Used Software and Platform.

Bibliography

[1] R. Mercier and M. Grelon, "Meiosis in plants: ten years of gene discovery," *Cytogenetic and Genome Research*, vol. 120, no. 3-4, pp. 281–290, 2008.

[2] M. A. Prusicki, M. Balboni, K. Sofroni, Y. Hamamura, and A. Schnittger, "Caught in the act: Live-cell imaging of plant meiosis," *Frontiers in Plant Science*, vol. 12, 2021. [Online]. Available: https://doi.org/10.3389/fpls.2021.718346

[3] M. A. Prusicki *et al.*, "Live cell imaging of meiosis in arabidopsis thaliana," *Elife*, vol. 8, 2019.

[4] E. M. Meyerowitz, "Arabidopsis, a useful weed," *Cell*, vol. 106, no. 6, pp. 647–650, 2001.

[5] T. A. G. Initiative, "Analysis of the genome sequence of the flowering plant arabidopsis thaliana," *Nature*, vol. 408, no. 6814, pp. 796–815, 2000.

[6] M. Koornneef and D. Meinke, "The development of arabidopsis as a model plant," *The Plant Journal*, vol. 61, no. 6, pp. 909–921, 2010.

[7] S. J. Clough and A. F. Bent, "Floral dip: a simplified method for agrobacterium-mediated transformation of arabidopsis thaliana," *The Plant Journal*, vol. 16, no. 6, pp. 735–743, 1998.

[8] C. Somerville and E. Meyerowitz, "The arabidopsis book," *The American Society of Plant Biologists*, 2004.

[9] S. Armstrong, "A time course for the analysis of meiotic progression in arabidopsis thaliana," *Methods in Molecular Biology*, vol. 990, p. 12, 2013.

[10] X. Yang, L. Yuan, and C. A. Makaroff, "Immunolocalization protocols for visualizing meiotic proteins in arabidopsis thaliana: method 3," *Methods in Molecular Biology*, vol. 990, pp. 109–118, 2013.

[11] C. Chen and E. F. Retzel, "Analyzing the meiotic transcriptome using isolated meiocytes of arabidopsis thaliana," *Methods in Molecular Biology*, vol. 990, pp. 203–213, 2013.

[12] S. Armstrong, "Spreading and fluorescence in situ hybridization of male and female meiocyte chromosomes from arabidopsis thaliana for cytogenetical analysis," *Methods in Molecular Biology*, vol. 990, pp. 3–11, 2013.

[13] J. Liu and L.-J. Qu, "Meiotic and mitotic cell cycle mutants involved in gametophyte development in arabidopsis," *Molecular Plant*, vol. 1, no. 4, pp. 564–574, 2008.

[14] E. Sanchez-Moran, K. Osman, J. Higgins, M. Pradillo, N. Cuñado, G. Jones, and F. Franklin, "Asy1 coordinates early events in the plant meiotic recombination pathway," *Cytogenet Genome Res*, vol. 120, no. 3-4, pp. 302–312, 2008.

[15] G. Jones, S. Armstrong, A. Caryl, and F. Franklin, "Meiotic chromosome synapsis and recombination in arabidopsis thaliana; an integration of cytological and molecular approaches," *Chromosome Res*, vol. 11, no. 3, pp. 205–215, 2003.

[16] A. De Muyt, L. Pereira, D. Vezon, L. Chelysheva, G. Gendrot, A. Chambon, S. Lainé-Choinard, G. Pelletier, R. Mercier, F. Nogué, and M. Grelon, "A high throughput genetic screen identifies new early meiotic recombination functions in arabidopsis thaliana," *PLoS Genet*, vol. 5, no. 9, p. e1000654, 2009.

[17] T. Naranjo and E. Corredor, "Nuclear architecture and chromosome dynamics in the search of the pairing partner in meiosis in plants," *Cytogenet Genome Res*, vol. 120, no. 3-4, pp. 320–330, 2008.

[18] K. Osman, J. Higgins, E. Sanchez-Moran, S. Armstrong, and F. Franklin, "Pathways to meiotic recombination in arabidopsis thaliana," *New Phytol*, vol. 190, no. 3, pp. 523–544, 2011.

[19] Z. Yu, K. Haage, V. E. Streit, A. Gierl, and R. A. T. Ruiz, "A large number of tetraploid arabidopsis thaliana lines, generated by a rapid strategy, reveal high stability of neo-tetraploids during consecutive generations," *Theoretical and Applied Genetics*, vol. 118, pp. 1107–1119, 2009.

[20] A. Pecinka, W. Fang, M. Rehmsmeier, A. A. Levy, and O. M. Scheid, "Polyploidization increases meiotic recombination frequency in arabidopsis," *BMC Biology*, vol. 9, no. 24, 2011.

[21] S. Liu, Y. Yang, F. Wei, J. Duan, J. Braynen, B. Tian, G. Cao, G. Shi, and J. Yuan, "Autopolyploidy leads to rapid genomic changes in arabidopsis thaliana," *Theory in Biosciences*, vol. 136, pp. 199–206, 2017.

[22] M. H. Jørgensen, D. Ehrich, R. Schmickl, M. A. Koch, and A. K. Brysting, "Interspecific and interploidal gene flow in central european arabidopsis (brassicaceae)," *BMC Evolutionary Biology*, vol. 11, no. 346, 2011.

[23] J. C. del Pozo and E. Ramirez-Parra, "Whole genome duplications in plants: an overview from arabidopsis," *Journal of Experimental Botany*, vol. 66, pp. 6991–7003, 2015.

[24] P. Y. Novikova, N. Hohmann, and Y. V. de Peer, "Polyploid arabidopsis species originated around recent glaciation maxima," *Current Opinion in Plant Biology*, vol. 42, pp. 8–15, 2018.

[25] K. M. Culligan, C. E. Robertson, J. Foreman, P. Doerner, and A. B. Britt, "Atr and atm play both distinct and additive roles in response to ionizing radiation," *The Plant Journal*, vol. 48, no. 6, pp. 947–961, 2006.

[26] A. N. Sakamoto, V. T. Lan, V. P. Puripunyavanich, Y. Hase, Y. Yokota, N. Shikazono, M. Nakagawa, I. Narumi, and A. Tanaka, "A uvb-hypersensitive mutant in arabidopsis thaliana is defective in the dna damage response," *The Plant Journal*, vol. 60, no. 3, pp. 509–517, 2009.

[27] L. Vespa, M. T. Couvillion, E. Spangler, and D. E. Shippen, "Atm and atr make distinct contributions to chromosome end protection and the maintenance of telomeric dna in arabidopsis," *Genes & Development*, vol. 19, no. 18, pp. 2111–2115, 2005.

[28] J. Zhao, X. Gui, Z. Ren, H. Fu, C. Yang, W. Wang, Q. Liu, M. Zhang, C. Wang, A. Schnittger, and B. Liu, "Atm-mediated double-strand break repair is required for meiotic genome stability at high temperature," *The Plant Journal*, vol. 114, no. 2, pp. 403–423, 2023.

[29] Y. Wang, Z. Cheng, P. Lu, L. Timofejeva, and H. Ma, "Molecular cell biology of male meiotic chromosomes and isolation of male meiocytes in arabidopsis thaliana," *Methods in Molecular Biology*, vol. 990, pp. 3–11, 2014.

[30] S. Armstrong, "Spreading and fluorescence in situ hybridization of male and female meiocyte chromosomes from arabidopsis thaliana for cytogenetical analysis," *Methods in Molecular Biology*, vol. 990, pp. 3–11, 2013.

[31] J. Li, J. Yuan, X. Zhao, X. Zhang, and M. Li, "A simple and rapid method for visualization of male meiotic chromosomes in arabidopsis thaliana," *Bioscience, Biotechnology, and Biochemistry*, vol. 78, pp. 1026–1028, 2014.

[32] J. De Jaeger-Braet, L. Krause, A. Buchholz, and A. Schnittger, "Heat stress reveals a specialized variant of the pachytene checkpoint in meiosis of Arabidopsis thaliana," *The Plant Cell*, vol. 34, no. 1, pp. 433–454, 10 2021. [Online]. Available: https://doi.org/10.1093/plcell/koab257

[33] C. Allier, L. Hervé, C. Paviolo, O. Mandula, O. Cioni, W. Pierré, F. Andriani, K. Padmanabhan, and S. Morales, "Cnn-based cell analysis: From image to quantitative representation," *Frontiers in Physics*, vol. 9, 2022. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fphy.2021.776805

[34] D. Molho, J. Ding, Z. Li, H. Wen, W. Tang, Y. Wang, J. Venegas, W. Jin, R. Liu, R. Su, P. Danaher, R. Yang, Y. L. Lei, Y. Xie, and J. Tang, "Deep learning in single-cell analysis," 2022.

[35] M. Flores, Z. Liu, T. Zhang, M. Hasib, Y. Chiu, Z. Ye, K. Paniagua, S. Jo, J. Zhang, S. Gao, Y. Jin, Y. Chen, and Y. Huang, "Deep learning tackles single-cell analysis-a survey of deep learning for scrna-seq analysis," *Briefings in Bioinformatics*, vol. 23, no. 1, p. bbab531, 2022.

[36] T. Athaya, R. Ripan, X. Li, and H. Hu, "Multimodal deep learning approaches for single-cell multi-omics data integration," *Briefings in Bioinformatics*, vol. 24, no. 5, p. bbad313, 2023.

[37] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *arXiv preprint arXiv:1409.1556*, 2014.

[38] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[39] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[40] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.

[41] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[42] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and <0.5mb model size," *arXiv preprint arXiv:1602.07360*, 2016.

[43] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.

[44] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.

[45] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International Conference on Machine Learning*, 2019, pp. 6105–6114.

[46] X. Qin, Z. Zhang, C. Huang, M. Dehghan, O. R. Zaiane, and M. Jagersand, "U^2-Net: Going deeper with nested u-structure for salient object detection," *Pattern Recognition*, vol. 106, p. 107404, 2020.

[47] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[48] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *2016 Fourth International Conference on 3D Vision (3DV)*. IEEE, 2016, pp. 565–571.

[49] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, B. Glocker, and D. Rueckert, "Attention u-net: Learning where to look for the pancreas," in *Proceedings of the 1st Conference on Medical Imaging with Deep Learning (MIDL 2018), Amsterdam, Netherlands*, 2018.

[50] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp. 3–11, 2018.

[51] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3d u-net: Learning dense volumetric segmentation from sparse annotation," in *International conference on medical image computing and computer-assisted intervention*. Springer, 2016, pp. 424–432.

[52] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[53] N. Abraham and N. M. Khan, "A novel focal tversky loss function with improved attention u-net for lesion segmentation," *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pp. 683–687, 2019.

[54] M. Tan, R. Pang, and Q. V. Le, "Efficientdet: Scalable and efficient object detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 10 781–10 790.

[55] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *European conference on computer vision*. Springer, 2016, pp. 21–37.

[56] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 580–587.

[57] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, 2015, pp. 91–99.

[58] X. Zhou, D. Wang, and P. Krähenbühl, "Objects as points," *arXiv preprint arXiv:1904.07850*, 2019.

[59] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1440–1448.

[60] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (voc) challenge," *International journal of computer vision*, vol. 88, no. 2, pp. 303–338, 2010.

[61] M. Maška, V. Ulman *et al.*, "The cell tracking challenge: 10 years of objective benchmarking," *Nature Methods*, vol. 20, no. 9, pp. 1010–1020, 2023. [Online]. Available: https://www.nature.com/articles/s41592-023-01879-y

[62] K. Bernardin and R. Stiefelhagen, "Evaluating multiple object tracking performance: The clear mot metrics," *EURASIP Journal on Image and Video Processing*, vol. 2008, no. 1, p. 246309, 2008. [Online]. Available: https://doi.org/10.1155/2008/246309

[63] M.-L. Zhang and Z.-H. Zhou, "A review on multi-label learning algorithms," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 8, pp. 1819–1837, 2014.

[64] G. Tsoumakas, I. Katakis, and I. Vlahavas, "Mining multi-label data," in *Data Mining and Knowledge Discovery Handbook*, O. Maimon and L. Rokach, Eds. Springer US, 2010, pp. 667–685.

[65] F. Herrera, F. Charte, A. J. Rivera, and M. J. del Jesus, "Multilabel classification: Problem analysis, metrics and techniques," *Journal of Computer Science and Technology*, vol. 31, pp. 161–175, 2016.

[66] G. Madjarov, D. Kocev, D. Gjorgjevikj, and S. Džeroski, "An extensive experimental comparison of methods for multi-label learning," *Pattern Recognition*, vol. 45, no. 9, pp. 3084–3104, 2012.

[67] Z.-H. Zhou and M.-L. Zhang, "Multi-instance multi-label learning with application to scene classification," in *Proceedings of the 19th International Conference on Neural Information Processing Systems*. MIT Press, 2006, pp. 1609–1616.

[68] J. Read, B. Pfahringer, G. Holmes, and E. Frank, "Classifier chains for multi-label classification," *Machine Learning*, vol. 85, no. 3, pp. 333–359, 2011.

[69] A. Elisseeff and J. Weston, "A kernel method for multi-labelled classification," in *Advances in Neural Information Processing Systems 14*, T. G. Dietterich, S. Becker, and Z. Ghahramani, Eds. MIT Press, 2002, pp. 681–687.

[70] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing & Management*, vol. 45, no. 4, pp. 427–437, 2009.

[71] M. Xue, J. Wang, L. Jiang, M. Wang, S. Wolfe, W. Pawlowski, Y. Wang, and Y. He, "The number of meiotic double-strand breaks influences crossover distribution in arabidopsis," *The Plant Cell*, vol. 30, pp. 2628–2638, 10 2018. [Online]. Available: https://doi.org/10.1105/tpc.18.00531

[72] G. Sidhu, C. Fang, M. Olson, M. Falque, O. Martin, and W. P. Pawlowski, "Recombination patterns in maize reveal limits to crossover homeostasis," *Proceedings of the National Academy of Sciences*, vol. 112, pp. 15 982–15 987, 2015. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.1514265112

[73] Ku, Ronceret, and Golubovskaya, "Dynamic localization of spo11-1 and conformational changes of meiotic axial elements during recombination initiation of maize meiosis," *PLOS Genetics*, vol. 16, pp. 1–30, 4 2020. [Online]. Available: https://doi.org/10.1371/journal.pgen.1007881

[74] J. Jing, T. Zhang, Y. Wang, Z. Cui, and Y. He, "Zmrad51c is essential for double-strand break repair and homologous recombination in maize meiosis," *International Journal of Molecular Sciences*, vol. 20, 2019. [Online]. Available: https://www.mdpi.com/1422-0067/20/21/5513

[75] Y. Wang and G. Copenhaver, "Meiotic recombination: Mixing it up in plants," *Annual Review of Plant Biology*, vol. 69, pp. 577–609, 2018. [Online]. Available: https://doi.org/10.1146/annurev-arplant-042817-040431

[76] Vignard, "The interplay of reca-related proteins and the mnd1-hop2 complex during meiosis in arabidopsis thaliana," *PLOS Genetics*, vol. 3, pp. 1–13, 10 2007. [Online]. Available: https://doi.org/10.1371/journal.pgen.0030176

[77] M. Serrentino and V. Borde, "The spatial regulation of meiotic recombination hotspots: Are all dsb hotspots crossover hotspots?" *Experimental Cell Research*, vol. 318, pp. 1347–1352, 2012, experimental Cell Research Special Review Issue: Chromosome Biology, 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0014482712001619

[78] C. Lambing and S. Heckmann, "Tackling plant meiosis: From model research to crop improvement," *Frontiers in Plant Science*, vol. 9, 2018. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpls.2018.00829

[79] Y. He, J. Wang, W. Qi, and R. Song, "Maize dek15 encodes the cohesin-loading complex subunit scc4 and is essential for chromosome segregation and kernel development," *The Plant Cell*, vol. 31, p. tpc.00921.2018, 1 2019.

[80] A. Sajid, E. Lalani, B. Chen, T. Hashimoto, D. Griffin, A. Bhartiya, G. Thompson, I. Robinson, and M. Yusuf, "Ultra-structural imaging provides 3d organization of 46 chromosomes of a human lymphocyte prophase nucleus," *International Journal of Molecular Sciences*, vol. 22, 2021. [Online]. Available: https://www.mdpi.com/1422-0067/22/11/5987

[81] A. Erilova, L. Brownfield, V. Exner, M. Rosa, D. Twell, O. M. Scheid, L. Hennig, and C. Köhler, "Imprinting of the polycomb group gene medea serves as a ploidy sensor in arabidopsis," *PLOS Genetics*, vol. 5, no. 9, pp. 1–14, 09 2009. [Online]. Available: https://doi.org/10.1371/journal.pgen.1000663

[82] N. De Storme and D. Geelen, "The Arabidopsis Mutant jason Produces Unreduced First Division Restitution Male Gametes through a Parallel/Fused Spindle Mechanism in Meiosis II," *Plant Physiology*, vol. 155, no. 3, pp. 1403–1415, 01 2011. [Online]. Available: https://doi.org/10.1104/pp.110.170415

[83] L. Brownfield, J. Yi, H. Jiang, E. A. Minina, D. Twell, and C. Köhler, "Organelles maintain spindle position in plant meiosis," *Nature Communications*, vol. 6, no. 1, p. 6492, 2015. [Online]. Available: https://doi.org/10.1038/ncomms7492

[84] R. Barnes, C. Lehman, and D. J. Mulla, "Priority-flood: An optimal depression-filling and watershed-labeling algorithm for digital elevation models," *CoRR*, vol. abs/1511.04463, 2015. [Online]. Available: http://arxiv.org/abs/1511.04463