



HOKKAIDO UNIVERSITY

Title	Development of a system for evaluating <i>Paeonia suffruticosa</i> pollen viability using deep learning techniques
Author(s)	ZHANG, Mengwei
Degree Grantor	北海道大学
Degree Name	博士(環境科学)
Dissertation Number	甲第16079号
Issue Date	2024-09-25
DOI	https://doi.org/10.14943/doctoral.k16079
Doc URL	https://hdl.handle.net/2115/97279
Type	doctoral thesis
File Information	ZHANG_Mengwei.pdf



**Development of a system for evaluating *Paeonia suffruticosa*
pollen viability using deep learning techniques**

(深層学習を用いた *Paeonia suffruticosa* の
花粉稔性の評価システムの開発)

ZHANG Mengwei

張 孟偉

Graduate School of Environmental Science, Hokkaido University

北海道大学大学院環境科学院

2024

Contents

Chapter 1: General introduction.....	1
1.1 <i>Paeonia suffruticosa</i>	1
1.2 Pollen viability and its role in plant hybrid breeding	1
1.3 Comparative overview of the methods for assessing pollen viability.....	2
1.4 Deep learning in plant science	3
1.5 Aims and flow	5
Chapter 2: Deep learning-based high-throughput detection of <i>in vitro</i> germination to assess pollen viability	7
2.1 Introduction	7
2.2 Materials and methods	8
2.2.1 Plant materials and image acquisition	8
2.2.2 Synthetic image generation	9
2.2.3 Image preparation and model training	9
2.2.4 Skeletonization of pollen tube mask	10
2.2.5 Evaluation of model performance	11
2.3 Results	11
2.3.1 Synthetic dataset improved the detection accuracy of the model	11
2.3.2 Evaluation of model training.....	12
2.3.3 Model performance on pollen grains and tube segmentation	12
2.3.4 Evaluation of model performance	13
2.3.5 Model practical evaluation	13

2.3.6 Model generalization test	14
2.3.7 Detection error.....	15
2.4 Discussion	16
2.5 Conclusion.....	19
Chapter 3: Morphology-based prediction of pollen germinability using deep-learning techniques in <i>Paeonia suffruticosa</i>.....	32
3.1 Introduction	32
3.2 Materials and methods	33
3.2.1 Plant material and culture conditions	33
3.2.2 Morphological analysis of hydrated pollen grains	34
3.2.3 Image acquisition and labeling of the deep-learning model	34
3.2.4 Determination of the optimal timing for capturing image M_I	35
3.2.5 Model training and evaluation.....	35
3.3 Results	36
3.3.1 Morphological analysis of pollen grains	36
3.3.2 Determination of the optimal time point of image capturing.....	37
3.3.3 Model evaluation.....	38
3.3.4 Model application and evaluation	38
3.4 Discussion	39
3.5 Conclusion.....	41
Chapter 4: General discussion	51
4.1 Advantages of the deep-learning technique in pollen viability assessment	51

4.2 Innovations of the approaches.....	51
4.3 Limitations and prospects.....	52
References	54
Summary	63
Acknowledgements	67
要約 (Summary in Japanese)	68

Abbreviations

AP	Average precision
AUC-PR	Area under the PR curve
FDA	Fluorescein diacetate
FDA-PI	FDA-propidium iodide
H ₂ DCFDA	2',7'-Dichlorodihydrofluorescein diacetate
IFC	Impedance flow cytometry
IoU	Intersection over Union
mAP	Mean average precision
PCA	Principal component analysis
RMSE	Root mean squared error
ROS	Reactive oxygen species
RPN	Region proposal network
R-CNN	Region convolutional neural network
YOLO	You only look once

Chapter 1: General introduction

1.1 *Paeonia suffruticosa*

Tree peony (*Paeonia suffruticosa*), a deciduous shrub of the family Paeoniaceae (Schoch et al., 2020), is renowned for its large and splendid flowers that come in various colors (Fig. 1.1), making it a popular ornamental plant in gardens and nurseries worldwide. New tree peony cultivars are developed using hybridization and selection of bud mutations (Cheng, 2007). The selection of seedlings obtained through cross-pollination is still the best way to breed new tree peony cultivars because these seedlings show better environmental adaptability than vegetatively propagated plants. This adaptability can be attributed to the genetic diversity of seed-propagated plants, which allows for greater resilience to varying environmental conditions and diseases (Crow, 1994; Roze, 2012).

1.2 Pollen viability and its role in plant hybrid breeding

An essential factor to consider hybrid breeding is pollen viability. Ensuring high pollen viability is crucial for achieving successful hybridization and developing robust cultivars during tree peony breeding. Pollen viability, a critical determinant of successful fertilization, reflects the ability of the pollen grains to deliver the sperm cells to embryo sac following compatible pollination (Shivanna et al., 1991, Dafni and Firmage, 2000). In rice (Khatun and Flowers, 1995) and potatoes (Trognitz, 1991), pollen viability and seed production were positively correlated with and highly viable pollen promoted seed formation. Additionally, pollen viability of the F1 generation can be a measure of F1 hybrid fertility, which can indicate post-zygotic reproductive isolation (Okuyama and Kato, 2009).

Pollen viability diminishes progressively after collection (Pacini et al., 1997, Boavida and McCormick, 2007, Brunet et al., 2019), highlighting the urgency to promptly assess pollen viability to ensure the optimal conditions for performing successful pollination.

1.3 Comparative overview of the methods for assessing pollen viability

Pollen viability can be assessed using numerous methods. The stigma provides species-specific nutrition supply and structural elements; therefore, *in vivo* or semi-*in vivo* tests are theoretically considered the most reliable approaches for testing the pollen viability of wheat and triticale (Esser, 1955; Fritz and Lukaszewski, 1989). However, the irregular growth of pollen tubes on the stigma makes it challenging to precisely identify pollen germination, particularly when different focal layers can be observed under the microscope (Dafni and Firmage, 2000, Impe et al., 2020). During Alexander staining (Alexander, 1969), the presence of protoplasm indicates viable pollen grain, which leads to an overestimation of pollen viability for some short-lived pollen grains, such as those of Johnsongrass (Burke et al., 2007), rye, barley, and maize (Impe et al., 2020). Fluorescein diacetate (FDA) staining can differentiate between aborted and non-aborted pollen by examining the integrity of the cytoplasmic membrane, which was highly correlated with pollen germination frequency (Abdul-Baki, 1992; Mazzeo et al., 2014). However, this integrity tends to overestimate pollen germination frequency (Wang et al., 2004; Dutta et al., 2013; Impe et al., 2020), particularly when applied to samples containing immature pollen (Helpson-Harrison et al., 1984). FDA-propidium iodide (FDA-PI) (Jones and Senft, 1985) double staining, an improvement over FDA staining, significantly improves the ability to classify between live and dead pollen (Dupl'áková

et al., 2016; Ascari et al., 2020). Furthermore, 2',7'-dichlorodihydrofluorescein diacetate (H₂DCFDA) staining involves using the non-cytotoxic reactive oxygen species (ROS) probe H₂DCFDA to predict pollen germination by measuring the ROS dynamics of pollen grains; this enables an accurate assessment of the viability of a large number of pollen grains when used in combination with flow cytometry (Luria et al., 2019; Rutley and Miller, 2020). Regarding non-staining methods, impedance flow cytometry (IFC) has also been used in pollen research (Heidmann et al., 2016; Impe et al., 2020); IFC estimates pollen viability and predicts pollen germination by detecting the electric properties in single pollen grains. However, the correlation between IFC viability scores and actual pollen germination capacity remains unclear (Luria et al., 2019). Go et al. (2019) assessed pollen viability based on the contour of pollen grains, but this approach is unsuitable for pollen that has undergone internal damage due to high temperature or long storage.

Therefore, although specific medium compositional modifications are required for various species and lines, *in vitro* pollen germination may be the most efficient method for determining *in vivo* pollen viability (Shivanna et al., 1991, Stone et al., 1995, Wang et al., 2004, Dutta et al., 2013).

1.4 Deep learning in plant science

Deep learning is a new method of artificial intelligence and uses artificial neural networks to discover patterns and connections in various datasets. The successful implementation of convolutional neural networks has significantly increased the precision of deep learning techniques in image recognition and trait prediction (Albawi et al., 2017).

Deep learning has been successfully applied in pollen research, most commonly in plant species identification using pollen images (Romero et al., 2020; Dunker et al., 2021;

Kubera et al., 2022). Additionally, deep learning has been used to determine the optimal developmental stage of pollen in eggplant haploid breeding (García-Fortea et al., 2020). As for research on the pollen activity assay, Tan et al. (2022) employed the You only look once (YOLO) model to assess pollen viability by classifying the pollen grains stained with 2,3,5-triphenyltetrazolium chloride; Yamazaki et al. (2023) trained a YOLO model for determining pollen germination rate. Compared to earlier algorithms of its kind, the advanced algorithm Mask Region-Based Convolutional Neural Network (R-CNN) is state-of-the-art for object detection and segmentation (He et al., 2017). Mask R-CNN recognizes objects in an image and simultaneously generates segmentation masks for each instance, which can then be used to analyze a series of morphological features in an object. A Mask R-CNN model was previously trained for crop seed phenotyping, which was used to improve important agronomic traits (Toda et al., 2020). Costa et al. (2021) trained a Mask R-CNN model for the area estimation of shuck, shell, and embryo of pecan nuts for yield prediction and growth phenology modeling. Warman et al. (2021) used Mask R-CNN to rapidly classify and quantify kernels in the maize ear, thereby providing large amounts of data for breeding. Mask R-CNN models are also used to analyze microscopic images. Xie et al. (2021) used the Mask R-CNN model combined with optical topology method to analyze the phenotype of maize stomata for QTL mapping. The Mask R-CNN model also facilitated high-throughput measurements of crossover frequency using fluorescent signal images of transgenic *Arabidopsis* pollen tetrads (Lim et al., 2020).

1.5 Aims and flow

This study aimed to facilitate pollen quality assessment and breeding for tree peony and other plant species. I employed two deep learning-based computer vision approaches for assessing pollen viability by analyzing microscopic images of pollen. Specifically, these approaches involved detecting or predicting pollen germination frequency and estimating pollen tube length.

Chapter 2 presents a method to detect pollen germination and estimate pollen tube length using the Mask R-CNN model. An improved Mask R-CNN model, which was trained on synthetic images, achieved high detection accuracy for tree peony pollen and demonstrated satisfactory performance in the generalization test.

Chapter 3 introduces a method to predict pollen germination frequency based on the morphology of hydrated tree peony pollen using the YOLOv7 model. Following model establishment, this method can even bypass *in vitro* pollen culture, thereby significantly shortening the detection time required. Furthermore, this model exhibits outstanding performance in handling a high density of pollen grains.

Chapter 4 concludes with a comprehensive discussion of the entire work.



Fig. 1.1. Flowers of the four tree peony cultivars used in this study.
(A) Cultivar ‘Shimadaijin’. (B) Cultivar ‘Shimanishiki’, (C) Cultivar ‘Houki’, (D)
Cultivar ‘Ten-i’.

Chapter 2: Deep learning-based high-throughput detection of *in vitro* germination to assess pollen viability

Chapter 2 is based on the publication in the Journal of Experimental Botany 74(21): 6551-6562 (2023) <https://doi.org/10.1093/jxb/erad315>. This paper is licensed under CC BY 4.0.

2.1 Introduction

Pollen germination frequency and tube length (Boavida and McCormick, 2007, Reichler et al., 2009, Jiang et al., 2015) are the key indicators of pollen viability. Pollen germination frequency is the proportion of pollen grains that successfully germinate, which can provide information about the overall health and fertility of pollen grains. Pollen germination frequency is affected by various abiotic stressors. A high germination frequency in pollen grains subjected to abiotic stresses, such as heat stress (Fahad et al., 2016; Luria et al., 2019) or water stress (Rang et al., 2011), indicates better pollen resistance, whereas a low pollen germination frequency indicates that the pollen grains have been severely damaged. Environmental factors, such as temperature and composition of the *in vitro* culture medium, affect pollen tube length. For instance, increasing the temperature from 24 to 36 °C significantly decreased pollen tube length, indicating a negative correlation between pollen tube length and temperature (Jiang et al., 2015, Poidevin et al., 2021). Additionally, substances, such as a microelement (Wang et al., 2003) or plant peptides (Covey et al., 2010), are added to the *in vitro* culture medium to assess their effect on germinability. In general, a good *in vitro* culture medium can induce longer pollen tubes with faster growth rates. Therefore, measuring the length of the pollen tube can help determine the conditions ideal for successful fertilization in a particular plant species.

However, using Mask R-CNN for estimating pollen germination frequency and pollen tube length has not been reported yet. To select pollen grains with high viability

or investigate the ideal conditions for pollen germination, pollen germination frequency and tube length must be determined based on numerous microscopic images. In this chapter, I present the rapid and accurate analysis of several microscopic images using the Mask R-CNN model to evaluate pollen viability.

2.2 Materials and methods

2.2.1 Plant materials and image acquisition

The pollen grains from tree peony (*Paeonia suffruticosa* ‘Shimanishiki’) were used for model training, testing, and practical assessment. The pollen grains of petunia (*Petunia hybrida*) and plantain lily (*Hosta plantaginea*) were used for the generalization test. Tree peonies and plantain lilies were grown under field conditions, whereas petunias were grown under greenhouse conditions. The field and greenhouse were located in the Field Science Center for Northern Biosphere, Hokkaido University, Japan. Pollen grain collection and culture were performed according to Hoshino et al. (2016). Mature anthers were sampled before dehiscence. The anthers were dried at 25 °C for 12 h and stored in a –20 °C fridge until further use. Pollen grains were vortex-mixed with 1 mL of liquid medium (Hirano and Hoshino, 2009) in a 2-mL centrifuge tube, and subsequently transferred to a 3.5-cm plastic petri dish for microscopic observation. The liquid medium contained 0.01% (w/v) H₃BO₃, 0.01% (w/v) CaCl₂, 0.0007% (w/v) KH₂PO₄, 0.01% (w/v) yeast extract, and 10% (w/v) sucrose at pH 5.8; the culture medium was sterilized via autoclaving before use.

All the images were captured using an inverted microscope (Axiovert 200; Carl Zeiss) with a 5× objective lens. According to the pre-experiment results, one anther per culture dish was recommended for tree peonies, so the density of pollen grains would not

exceed 40 per image. For model training, testing, and generalization testing, images were randomly captured within 0.5–3 h after pollen culture. A practical evaluation experiment was performed to test the practical application of the proposed model; 20 images were captured at 0.5, 1, 2, and 3 h after pollen culture. In the figures, some microscopic images are zoomed in to highlight the details. Scale bars have been inserted to indicate the size of the samples.

2.2.2 Synthetic image generation

First, independent individual germinated pollen was outlined from the original microscopic images using the Polygonal Lasso Tool in Photoshop 2020 and saved as a newly created layer. In total, 100 different single pollen images were selected to create a pollen tube image pool. Second, 20 microscopic images, excluding pollen grains with a fixed size of 692×520 pixels, were prepared as the background image pool. Third, pollen images from the pollen tube image pool were randomly selected and manually placed on the background images from the background image pool to generate synthetic images for simulating the real scenario of pollen tube crossings (Fig. 2.1). Finally, 100 sets of synthetic images were generated similarly.

2.2.3 Image preparation and model training

Fig. 2.2 shows the process and workflow for preparing the training data of the proposed model. In total, 500 original and 100 synthetic images were prepared with a size of 692×520 pixels at 300 dpi. A total of 600 images were manually labeled using Make Sense (<https://www.makesense.ai/>) and subjected to data augmentation (50% probability of horizontal flipping; 50% probability of vertical flipping; random brightness adjustment

of between -20 and 0 %; random Gaussian blur between 0 and 0.5 pixels; salt and pepper noise applied to 1 percent of the pixels) to generate 1800 images, of which 1620 contained 14727 and 8863 instances of pollen grains and tubes, respectively, and were used as the training data; the other 180 images were used as the validation data. The test dataset comprised 120 original microscopic images with an image size of 1280×960 pixels at 300 dpi; these were not used for model training or validation.

The Mask R-CNN algorithm was selected to perform instance segmentation of pollen grains and pollen tubes. The detailed settings of the neural network frame are summarized in Table 2.1. Validation loss, mean average precision (mAP), and accuracy were calculated to assess the model quality. Validation loss was measured after each epoch as a metric to assess the performance of the deep-learning model on the validation set. Average precision (AP) was used to evaluate the classification performance, and mAP was estimated as the average of AP over all the detected classes, including pollen grains and tubes. Accuracy was a measure of the fraction of correctly predicted pollen grains and tubes.

2.2.4 Skeletonization of pollen tube mask

When the model is initiated, the microscopic images are input into the trained model for detection. After obtaining the raw output with bounding boxes and mask regions, the pollen tube mask is extracted and transformed into a binary image. The skeletonization is performed based on this binary image (Fig. 2.2). The medial axis algorithm (Marie et al., 2016) was selected for skeletonization. Skeletonization was performed to extract a one-pixel-wide shape feature representing the general form of a pollen tube, which could then be used to precisely measure the real length of the pollen tube using a scale bar.

2.2.5 Evaluation of model performance

Pollen germination frequency was calculated as follows:

$$\text{Pollen germination frequency} = \frac{\text{No. of pollen tubes}}{\text{No. of pollen grains}}$$

The average length of pollen tubes was calculated as follows:

$$\text{Average length of pollen tube} = \frac{\text{Total length of pollen tubes}}{\text{No. of pollen tubes}}$$

The ground truth of pollen tube length was manually measured using ImageJ 2.9.0 (Abràmoff et al., 2004). Data processing and charting were performed using Python and GraphPad Prism 8.

2.3 Results

2.3.1 Synthetic dataset improved the detection accuracy of the model

The trained pollen detection model was used to segment the pollen grains and tubes to extract relevant traits. When the input image contained cross-pollen tubes, the model trained without synthetic images exhibited several errors during detection (Fig. 2.3); for example, only a single pollen tube was detected among crossed pollen tubes or evident pollen tube segmentation errors. However, the segmentation quality improved considerably after adding synthetic images to the training dataset and maintaining the original parameters (Fig. 2.3). The updated model allowed for better identification and segmentation of crossed pollen tubes, thereby reducing the errors in the calculations of pollen germination frequency and tube length.

2.3.2 Evaluation of model training

mAP and validation loss were calculated to evaluate the performance of the proposed model (Fig. 2.4). The volatility of the validation loss of the proposed model decreased over 200 epochs, with the lowest validation loss reducing to 0.218 and mAP stabilizing around 0.95 after approximately 150 epochs. Based on these results, the model with a low validation loss and high mAP was chosen for subsequent analysis. Model evaluation based on the validation data is shown in Table 2.2. At the Intersection over Union (IoU) threshold of 50%, the mAP was 0.949. For higher IoU thresholds, the model still resulted commendable scores of 0.820 and 0.633 for mAP₇₀ and mAP@ [.5:.95], respectively. The model achieved accuracies of 0.997 for pollen grains and 0.903 for pollen tubes at an IoU threshold of 50%.

2.3.3 Model performance on pollen grain and tube segmentation

The performance of the proposed model is illustrated in Fig. 2.5. The input image shown in Fig. 2.5A was obtained from the test dataset captured using a 5× objective lens. The input image was processed by the proposed model for approximately 26 s and then outputted first as a raw output (Fig. 2.5B), subsequently as a binary image of pollen tube masks (Fig. 2.5C), and finally as a skeletonized image of pollen tubes (Fig. 2.5D).

In the raw output, almost all the pollen grains and tubes were segmented correctly, and the generated masks covered the target objects accurately. As expected from the circumstances of each image, the model correctly identified not only the large and round mature but also the small and oval immature pollen grains. In particular, for pollen grain clusters, which are pollen grains that are tightly packed together, the model could segment each individual precisely. Simultaneously, the predicted number of pollen grains and

tubes was automatically counted and outputted. Based on the raw output, binary masks of pollen tubes were extracted (Fig. 2.5C) and subjected to skeletonization; the binary masks were skeletonized to 1-pixel-wide line segments (Fig. 2.5D). After counting the total pixels, the pollen tube length was calculated according to the scale. Based on the image size of the test dataset, the scale used in this experiment was 760 pixels, which is equal to 1 mm.

2.3.4 Evaluation of model performance

Fig. 2.6 shows four scatter plots comparing the ground truth and model predictions for the four indexes (number of pollen grains, number of pollen tubes, pollen germination frequency, and average length of pollen tubes) calculated from the 120 test images. The R-squared of the fitted equation was >0.9 for all the four indexes, and that for three indexes was close to 1, indicating that the model was well-fitted and robust. The slope values of the fitted equations were very close to 1 for the number of pollen grains (Fig. 2.6A), number of pollen tubes (Fig. 2.6B), and pollen germination frequency (Fig. 2.6C), indicating that the proposed model achieved high accuracy, as expected. Notably, while the model fit was good in the case of average pollen tube length prediction (Fig. 2.6D), the slope value reached only 0.797, indicating a poorer prediction of the average length compared to the ground truth.

2.3.5 Model practical evaluation

A practical evaluation was performed to verify the effectiveness of the model for practical applications and detecting pollen tubes at different developmental stages. Images of developing pollen tubes were captured at four time points within 3 h of tree peony pollen

culture. Pollen germination frequency and average pollen tube length were calculated via manual counting and model prediction, respectively. The model maintained a high level of accuracy for the detection of pollen grains and tubes. The model detected pollen tubes of different lengths, from very short to long (Fig. 2.7), and with different curvatures (Fig. 2.8).

The pollen germination frequency increased from 57.5% at 0.5 h to 70.2% at 2 h and did not significantly increase further (Table 2.3). No significant difference was observed between the model predictions and ground truth at any of the four time points (Fig. 2.9A). The average pollen tube length increased consistently from 45.8 μm at 0.5 h to 345.9 μm at 3 h (Table 2.3). Although model prediction did not differ significantly from the ground truth at 0.5 h, significant differences were observed at the other three time points (Fig. 2.9B).

2.3.6 Model generalization test

Two horticultural plants, petunia and plantain lily, were selected for the model generalization test. The morphology of the pollen grains and tubes in both the species differs considerably from those of tree peony. The pollen grains of petunia were smaller and lighter in color than those of tree peony, and the pollen tubes were more likely to grow curved compared with those of the plantain lily, whose pollen grains were larger and had a dark brown appearance. Additionally, the pollen tubes of plantain lily were thicker and longer than those of tree peony. Surprisingly, the original output was excellent, with pollen grains and tubes clearly segmented in petunia (Fig. 2.10A) and plantain lily (Fig. 2.10B). A comparison of the detected results with the ground truth showed that the

model achieved the expected results for petunia and plantain lily, as illustrated in Figs. 2.10C and 2.10D, respectively.

2.3.7 Detection error

The model accurately estimated the pollen length when the pollen tube was not too long; however, with increase in pollen tube length, the errors between the estimated length and ground truth increased gradually (Fig. 2.9B). These errors can be caused due to multiple reasons, including partial detection, one pollen tube being detected as two, and length loss during skeletonization (Fig. 2.11). During partial detection (Fig. 2.11A), the possibility of errors increased when the pollen tube was long or when multiple pollen tubes crossed, resulting in a decrease in the estimated length. When the density of pollen grains was too high, the overlap between pollen grains and tubes increased the likelihood of errors, where one pollen tube was mistakenly detected as two (Fig. 2.11B). Thus, the overestimation of the number of pollen tubes resulted in the underestimation of the average pollen tube length. In general, a small segment of length at both ends of the pollen tube mask was lost because of the thinning effect during skeletonization (Saeed et al., 2010); thus, the combined image shows a slightly shorter skeletonized line than the binary image of pollen tube masks (Fig. 2.11C). Although the impact of this error is relatively small compared to the first two cases, it still causes observable errors at higher pollen tube densities.

2.4 Discussion

In this study, a Mask R-CNN model was developed using microscopic images of pollen from tree peonies to rapidly detect pollen germination frequency and tube length. Including synthetic images in the training set significantly increased the recognition accuracy of the model for crossed pollen tubes. Moreover, the model demonstrated strong adaptability and generalizability during testing with petunia and plantain lily samples.

The initial model performed poorly when segmenting crossed pollen tubes because of the small sample size during training. However, obtaining images of crossed pollen tubes to be used in the training dataset is time consuming. Therefore, a method to generate synthetic images was employed (Figs. 2.1 and 2.3). This enabled the acquisition of a sufficient number of images of crossed pollen tubes for training within one day, negating the need for additional data collection. Synthetic image generation is a fast means of acquiring a large number of training images while reducing the workload of manual annotation. It can also be used to simulate images that are difficult to collect in experimental settings, thereby rapidly enriching training datasets. Tian et al. (2020) enriched the training dataset using synthetic images that combined the background and apple flower images based on the growth traits of apple flowers, significantly improving the performance of the model in complex orchard scenes. In our study, synthetic images were used to enrich the training dataset, which significantly improved the recognition accuracy of the model for crossed pollen tubes.

To calculate the pollen germination frequency, instance segmentation of pollen grains and tubes was performed instead of categorizing individual pollen grains as germinated or non-germinated pollen grains (Fig. 2.5B). Subsequently, the number of each kind was counted to calculate the pollen germination frequency (Figs. 2.6A and

2.6B). The results showed that our model could accurately calculate the germination frequency, regardless of pollen tube length (Figs. 2.6C and 2.9A), and that instance segmentation of the pollen grains and tubes also facilitated the separate analyses of pollen grain or tube morphology. Pollen tube masks generated via instance segmentation were used to estimate pollen tube length (Fig. 2.5C). In previous studies, masks were commonly used for fitting analysis; the features of objects with regular shapes can be extracted using algorithms. For example, Toda et al. (2020) fitted the mask of hundreds of barley seeds into a general barley seed shape, and Xie et al. (2021) fitted the mask of stomata into ellipses. However, it is difficult to fit the pollen tube image data using any specific algorithm because they lack a fixed shape. Therefore, in this study, the length of the pollen tube was measured by skeletonizing the pollen tube mask (Fig. 2.5D), a method that has been employed in engineering to measure crack length (Yang et al., 2018; Dang et al., 2022). Overall, the performance of our model in calculating the pollen germination frequency and average pollen tube length was acceptable (Figs. 2.6D and 2.9B). However, the pollen grain density should not be so high to avoid the excessive occurrence of overlapping pollen grains and crossed pollen tubes.

Although all the images in the training dataset were sourced from tree peony, the model also demonstrated good performance when applied directly to two other horticultural plants, petunia and plantain lily, despite the differences in their pollen grain size, color, and tube morphology (Fig. 2.10). The results of the generalizability test suggest that our model is adaptable to pollen grains and tubes within a certain range of similarity. Transfer learning, a method that uses the architecture and weights of a pre-trained neural network, can be used to train a model with a small amount of training data to rapidly improve its performance on a new task without reconstructing the model (Kaya

et al., 2019; Neyshabur et al., 2020). In future studies, the use of transfer learning is recommended for species with excessive morphological differences in pollen grains and tubes or for microscopic images with different backgrounds and magnifications. Therefore, the proposed model has strong application potential in other plant species or under various microscopy conditions. In previous studies, solid media were commonly used for *in vitro* pollen culture. Pollen grains were unevenly distributed and accumulated easily on the solid medium (Boavida and McCormick, 2007; Impe et al., 2020), making it difficult to determine the pollen germination status under a microscope. In the present study, I used a specific liquid medium developed in our laboratory for *in vitro* pollen culture (Hirano and Hoshino, 2009), modified based on a previous method (Hoshino et al., 2016). First, the pollen grains were vortex-mixed with the culture medium in a centrifuge tube before being transferred into a petri dish to provide uniform environmental conditions. In these conditions, pollen grains settled at the bottom of the dish. Second, I reduced the volume of the culture medium to 1 mL and used a 5× objective to minimize the loss of field of view due to out-of-focus. In a liquid medium, pollen grains distribute evenly and do not overlap with each other when used at an appropriate pollen concentration, making it easy to observe and positively impact the accuracy of model identification. More importantly, pollen cultured in a liquid medium can be isolated for subsequent analysis, as demonstrated by Hirano and Hoshino (2009, 2010), who used a liquid medium to examine the changes in the nuclear phase during pollen tube development. Uchida et al. (2012) suggested that pollen tubes cultivated in a liquid medium can be easily recovered and used as materials for molecular biological analyses, such as proteome analysis. However, these advantages are difficult to achieve using solid media.

2.5 Conclusion

To rapidly evaluate pollen viability, a deep-learning model was trained based on Mask R-CNN to estimate pollen germination frequency and tube length. This model has the application and generalization potential in other plant species. Furthermore, the model can be used in various scenarios, including determining the optimal temperature for pollen germination, optimizing pollen culture media, and evaluating stored pollen viability.

Table 2.1. Settings for the Mask R-CNN model.

Network	Parameter
Backbone	ResNet101
Optimizer	Stochastic Gradient Descent (SGD)
Learning rate	1.0E-03
Momentum	0.9
Decay	1.0E-06
Batch size	8
Epochs	200

This table is from “Deep learning-based high-throughput detection of in vitro germination to assess pollen viability from microscopic images” (DOI: 10.1093/jxb/erad315) © 2023 by Mengwei Zhang, Jianxiang Zhao, Yoichiro Hoshino licensed under CC BY 4.0.

Table 2.2. Model evaluation based on the validation data.

	Mask region metrics			Accuracy	
	mAP ₅₀	mAP ₇₀	mAP@ [.5:.95]	pollen grain	pollen tube
Values	0.949	0.820	0.633	0.997	0.903

Mean average precision (mAP) at an IoU threshold of 50% (mAP₅₀) and 70% (mAP₇₀), and the average mAP from IoU 50–95% with a step size of 5% (AP@ [.5:.95]). Accuracy values at the IoU threshold of 50% are shown.

This table is from “Deep learning-based high-throughput detection of in vitro germination to assess pollen viability from microscopic images” (DOI: 10.1093/jxb/erad315) © 2023 by Mengwei Zhang, Jianxiang Zhao, Yoichiro Hoshino licensed under CC BY 4.0.

Table 2.3. Data of the practical evaluation experiment.

		0.5 h	1 h	2 h	3 h
Pollen germination frequency	Ground truth	57.5±13.6 ^b	65.5±10.1 ^{ab}	70.2±8.8 ^a	70.6±12.2 ^a
	Detection	57.9±10.7 ^b	66.9±10.7 ^{ab}	70.9±8.7 ^a	73.2±13.9 ^a
Average length of pollen tubes	Ground truth	45.8±15.9 ^c	119.9±20.9 ^d	232.9±35.5 ^c	345.9±66.0 ^a
	Detection	41.8±13.2 ^c	106.5±19.2 ^d	203.1±32.0 ^c	284.5±54.9 ^b

Values are presented as the mean ± SD of twenty replicates. Different letters indicate significant differences at $P < 0.05$ (Student’s *t*-test).

This table is from “Deep learning-based high-throughput detection of in vitro germination to assess pollen viability from microscopic images” (DOI: 10.1093/jxb/erad315) © 2023 by Mengwei Zhang, Jianxiang Zhao, Yoichiro Hoshino licensed under CC BY 4.0.

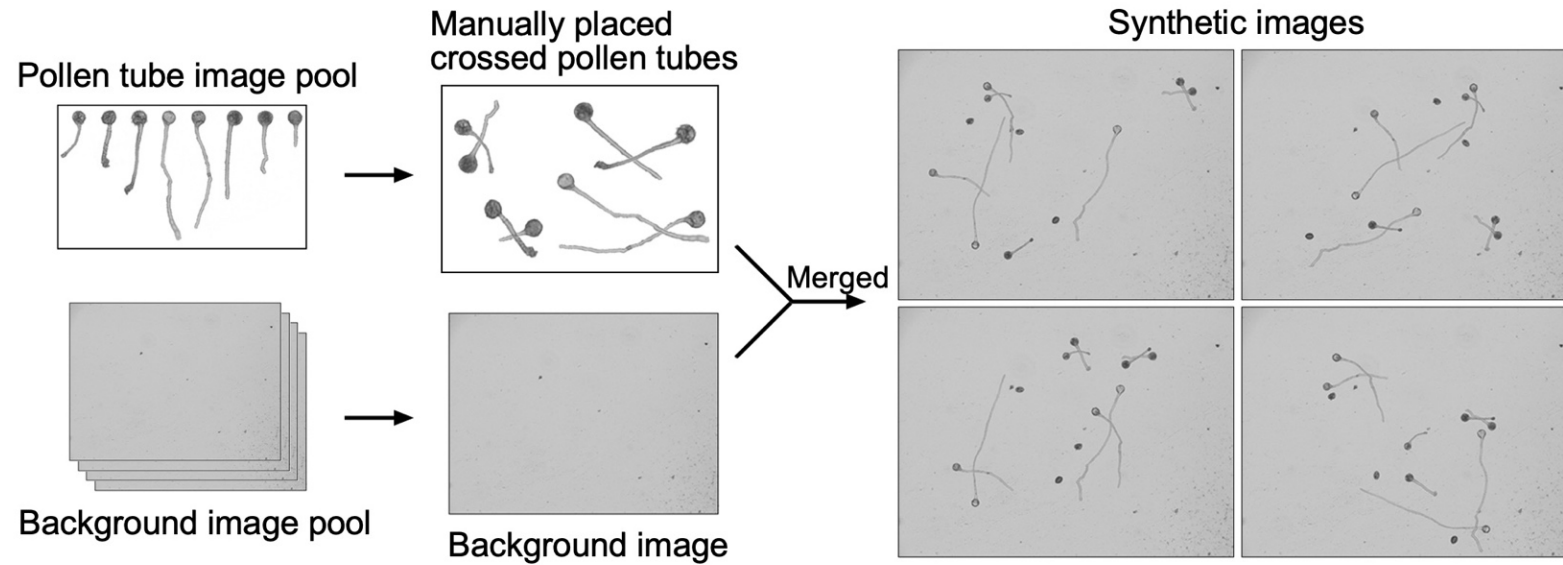
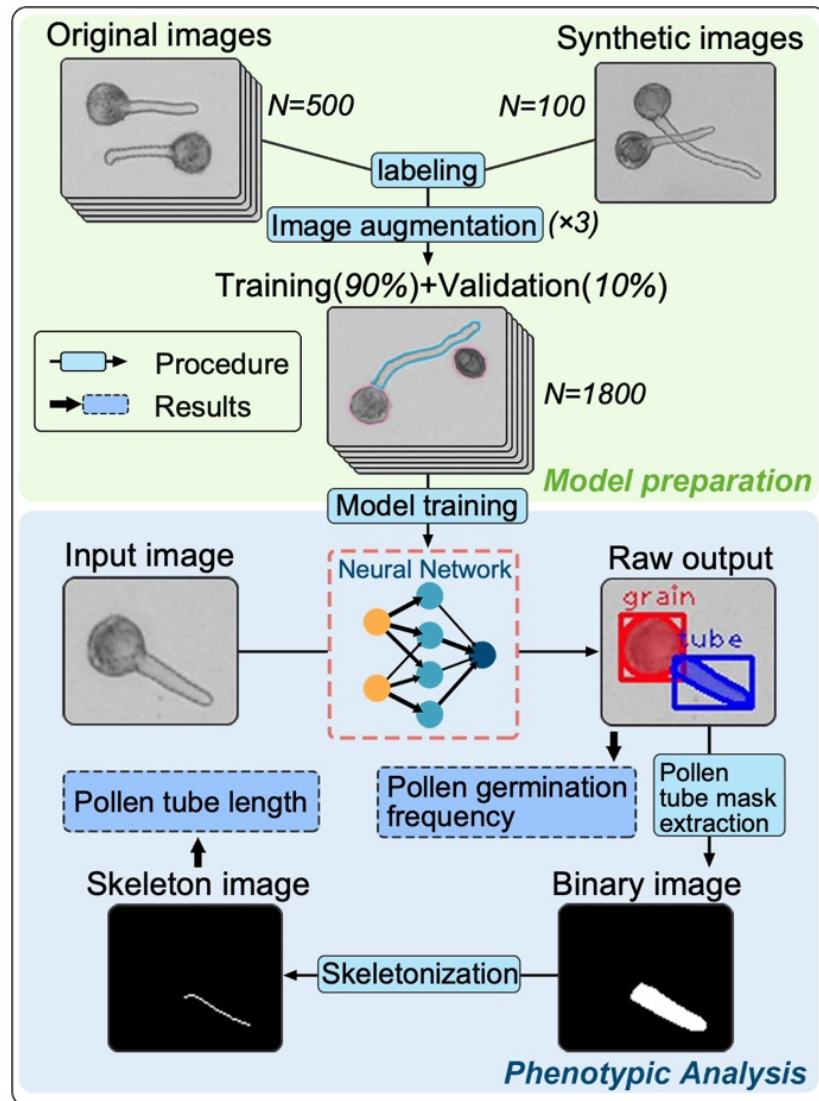


Fig. 2.1. Creating of synthetic images.

First, individual germinated pollen grains were isolated from the original microscopic image using the Polygonal Lasso Tool in Photoshop 2020 and saved as new layers. A total of 100 different single pollen images were selected to create a pollen tube image pool. Second, 20 microscopic images, each 692×520 pixels and excluding the pollen grains, were prepared as the background image pool. Finally, pollen images from the pollen tube image pool were randomly selected and manually placed onto the background images from the background image pool to generate synthetic images, simulating the real scenario of pollen tube crossings.

This figure is from “Deep learning-based high-throughput detection of in vitro germination to assess pollen viability from microscopic images” (DOI: 10.1093/jxb/erad315) © 2023 by Mengwei Zhang, Jianxiang Zhao, Yoichiro Hoshino licensed under CC BY 4.0.

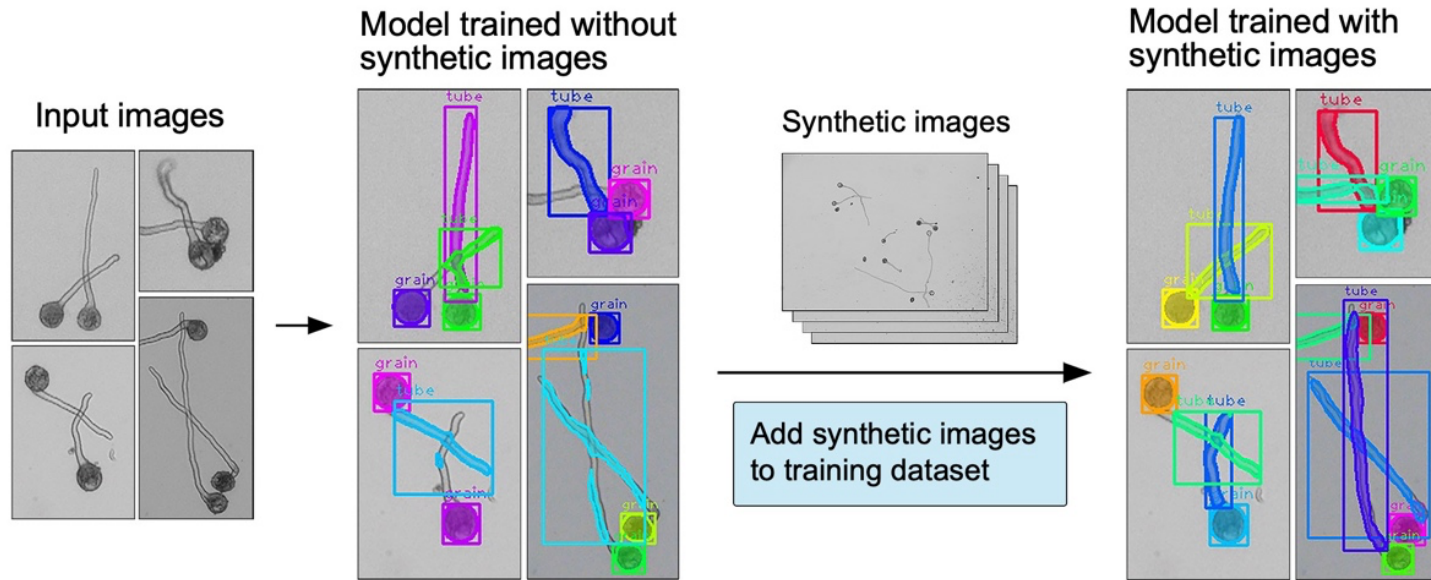


12

13 **Fig. 2.2.** Workflow and the working principle of the proposed Mask R-CNN model.14 *This figure is from "Deep learning-based high-throughput detection of in vitro germination to*15 *assess pollen viability from microscopic images" (DOI: 10.1093/jxb/erad315) © 2023 by*16 *Mengwei Zhang, Jianxiang Zhao, Yoichiro Hoshino licensed under CC BY 4.0.*

17

18
19
20



21

22 **Fig. 2.3.** Comparison of the detection performances of the models with and without synthetic tree peony pollen microscopic images.

23 Adding synthetic images to the training dataset while maintaining the original parameters enables the updated model to better recognize and
24 segment crossed pollen tubes. This reduces errors in calculating pollen germination frequency and tube length.

25 *This figure is from “Deep learning-based high-throughput detection of in vitro germination to assess pollen viability from microscopic images” (DOI:*
26 *10.1093/jxb/erad315) © 2023 by Mengwei Zhang, Jianxiang Zhao, Yoichiro Hoshino licensed under CC BY 4.0.*

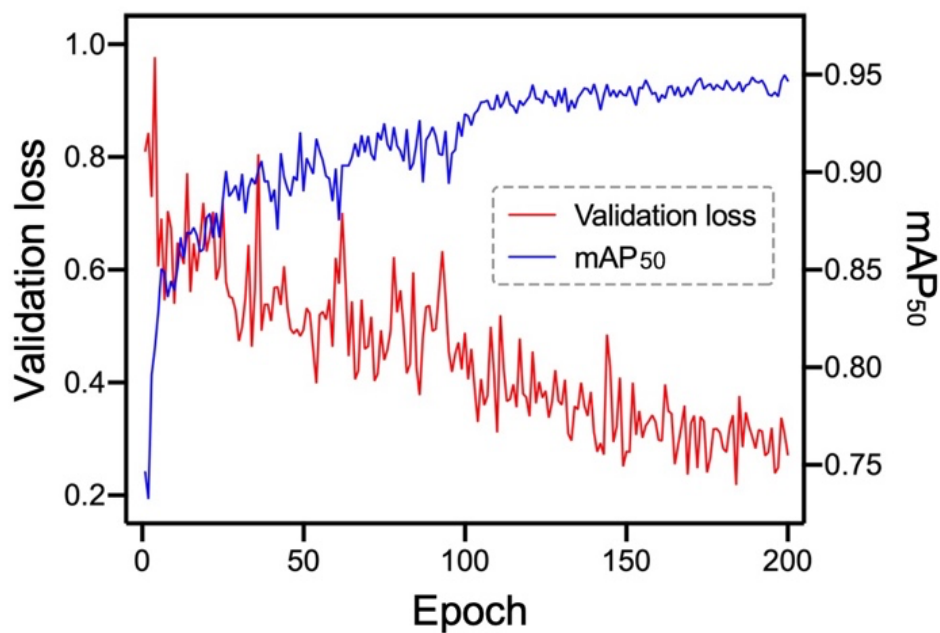


Fig. 2.4. Validation loss and mean average precision (mAP) at the IoU threshold of 50% (mAP₅₀) of models based on 180 validation images during 200 epochs.

This figure is from “Deep learning-based high-throughput detection of in vitro germination to assess pollen viability from microscopic images” (DOI: 10.1093/jxb/erad315) © 2023 by Mengwei Zhang, Jianxiang Zhao, Yoichiro Hoshino licensed under CC BY 4.0.

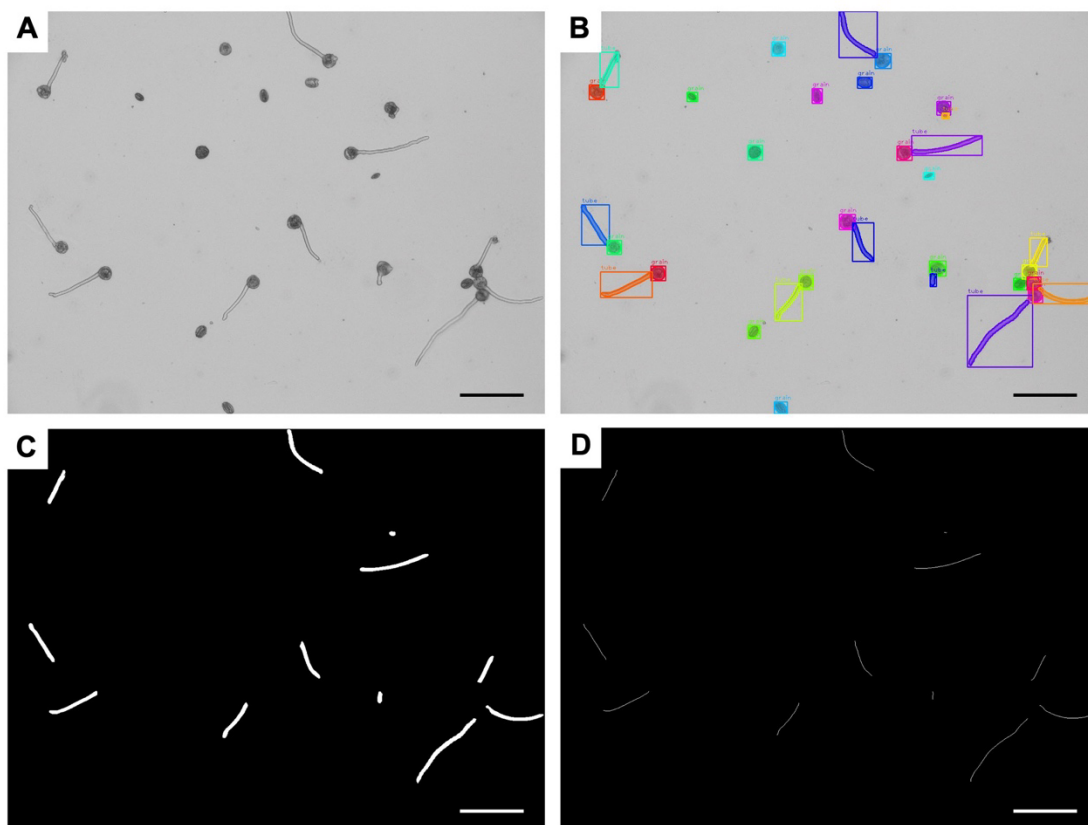


Fig. 2.5. Output of the designed model compared to the input image.

(A) Input image, (B) raw output, (C) binary mask of pollen tubes, and (D) skeleton of pollen tubes. Scale bars = 200 μm .

This figure is from “Deep learning-based high-throughput detection of in vitro germination to assess pollen viability from microscopic images” (DOI: 10.1093/jxb/erad315) © 2023 by Mengwei Zhang, Jianxiang Zhao, Yoichiro Hoshino licensed under CC BY 4.0.

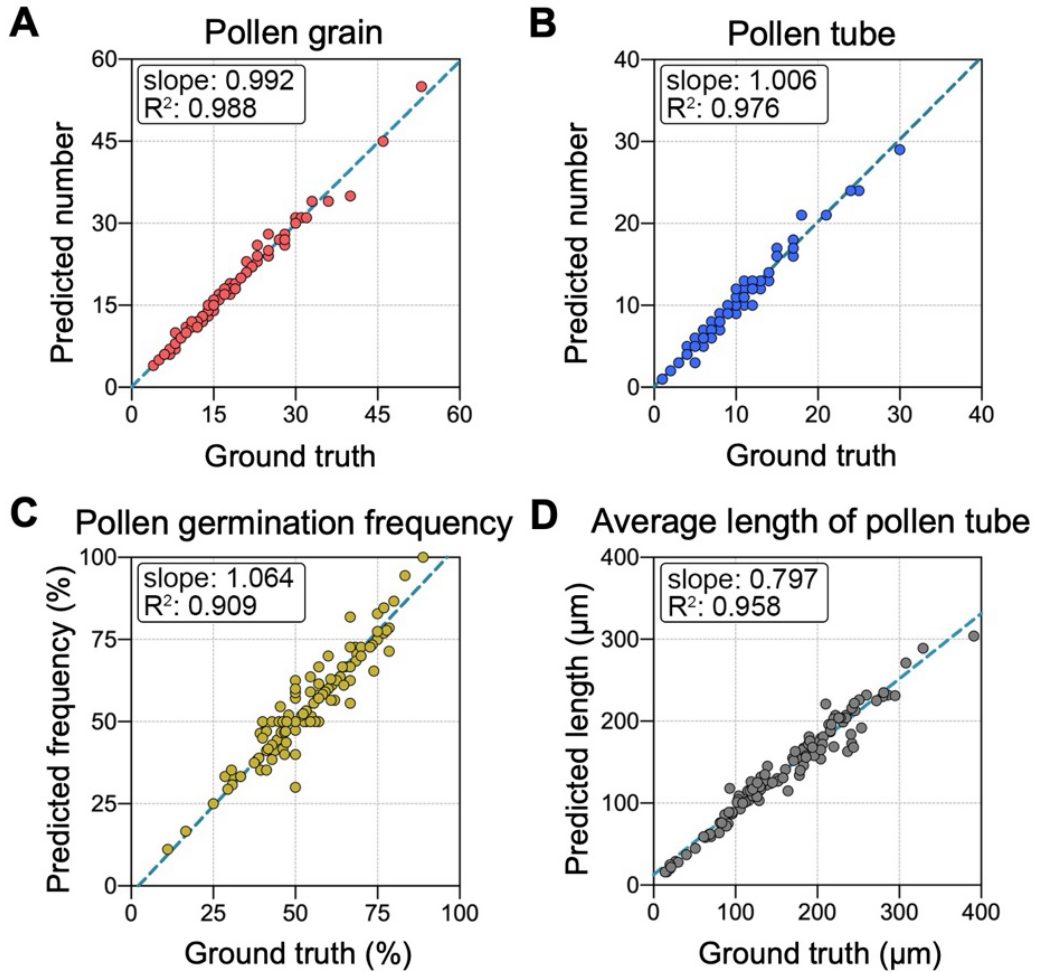


Fig. 2.6. Scatter plots of the results compared to ground truth for the four pollen germination indexes in the test dataset.

(A) No. of pollen grains, (B) no. of pollen tubes, (C) pollen germination frequency, and (D) average length of pollen tube. Each dot represents an original microscopic image. The dark green dotted lines represent the linear regression equations.

This figure is from “Deep learning-based high-throughput detection of in vitro germination to assess pollen viability from microscopic images” (DOI: 10.1093/jxb/erad315) © 2023 by Mengwei Zhang, Jianxiang Zhao, Yoichiro Hoshino licensed under CC BY 4.0.

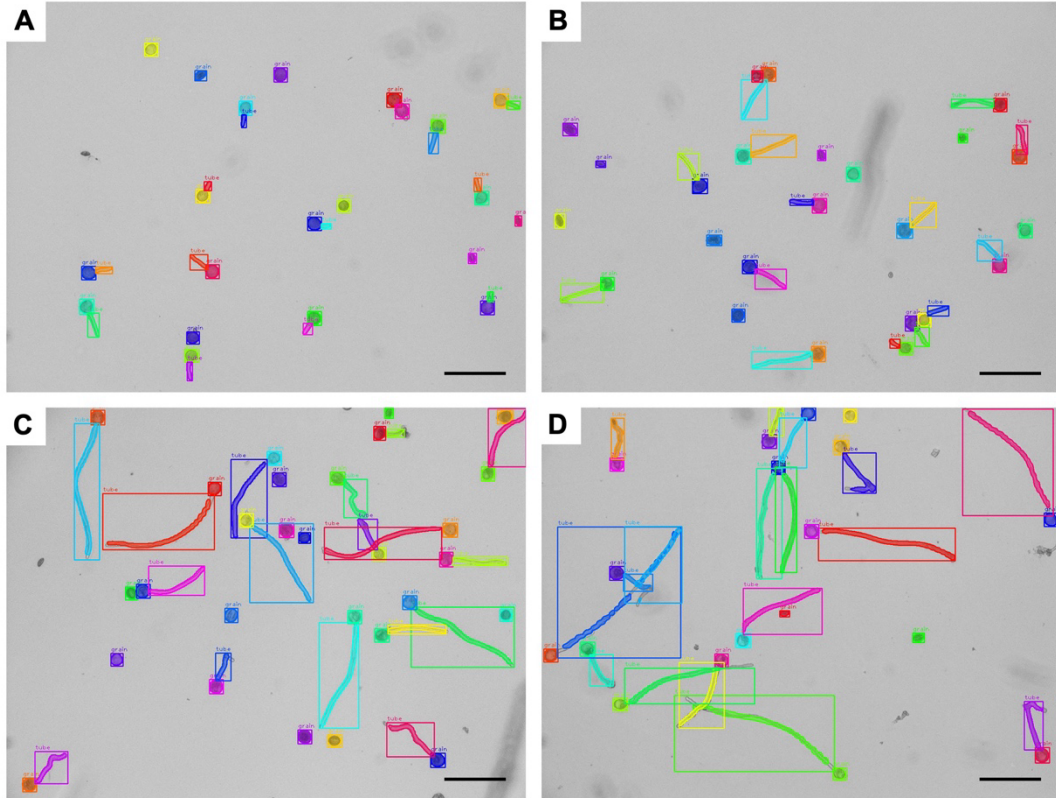


Fig. 2.7. Raw output of the proposed model for tree peony pollen microscopic images.

(A) Some pollen grains started to germinate at 0.5 h after culture. (B) The pollen tubes persistently elongated at 1 h after culture. (C) Some pollen tubes were long and overlapped each other at 2 h after culture. (D) Longer pollen tubes produced a more complex crossover situation at 3 h after culture. All the pollen grains and tubes with different lengths and curvatures were detected with high accuracy at all the time points.

Scale bars = 200 μm .

This figure is from “Deep learning-based high-throughput detection of in vitro germination to assess pollen viability from microscopic images” (DOI: 10.1093/jxb/erad315) © 2023 by Mengwei Zhang, Jianxiang Zhao, Yoichiro Hoshino licensed under CC BY 4.0.

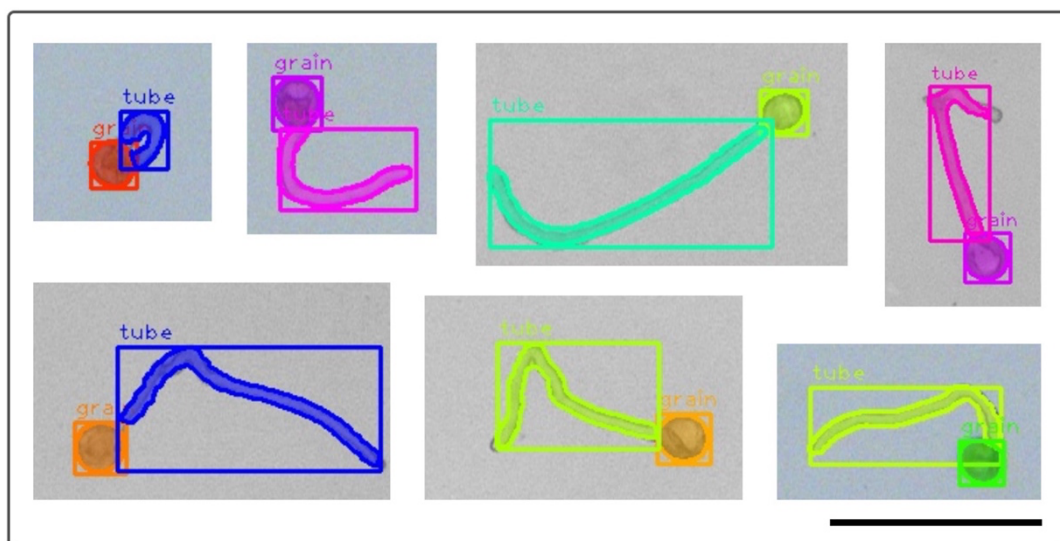


Fig. 2.8. Output of the model detecting pollen tubes with different curvatures.

Scale bars = 200 μm .

This figure is from “Deep learning-based high-throughput detection of in vitro germination to assess pollen viability from microscopic images” (DOI: 10.1093/jxb/erad315) © 2023 by Mengwei Zhang, Jianxiang Zhao, Yoichiro Hoshino licensed under CC BY 4.0.

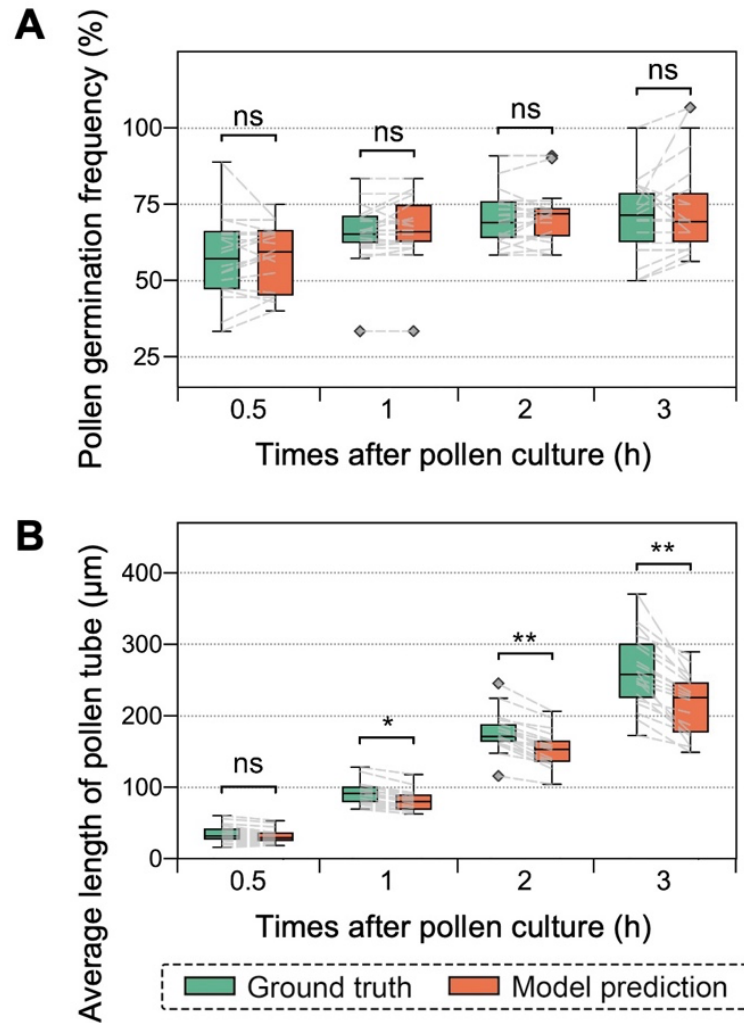


Fig. 2.9. Practical evaluation of tree peony pollen germination at 0.5, 1, 2, and 3 h after pollen culture.

(A) Pollen germination frequency. (B) Average length of the pollen tube. Significant differences were analyzed using a t -test for independent samples and are indicated using different notations plotted above the box plots. *, $P < 0.05$; **, $P < 0.01$; ns, not significant. Diamond symbols represent outliers. Dashed lines link paired the same images.

This figure is from “Deep learning-based high-throughput detection of in vitro germination to assess pollen viability from microscopic images” (DOI: 10.1093/jxb/erad315) © 2023 by Mengwei Zhang, Jianxiang Zhao, Yoichiro Hoshino licensed under CC BY 4.0.

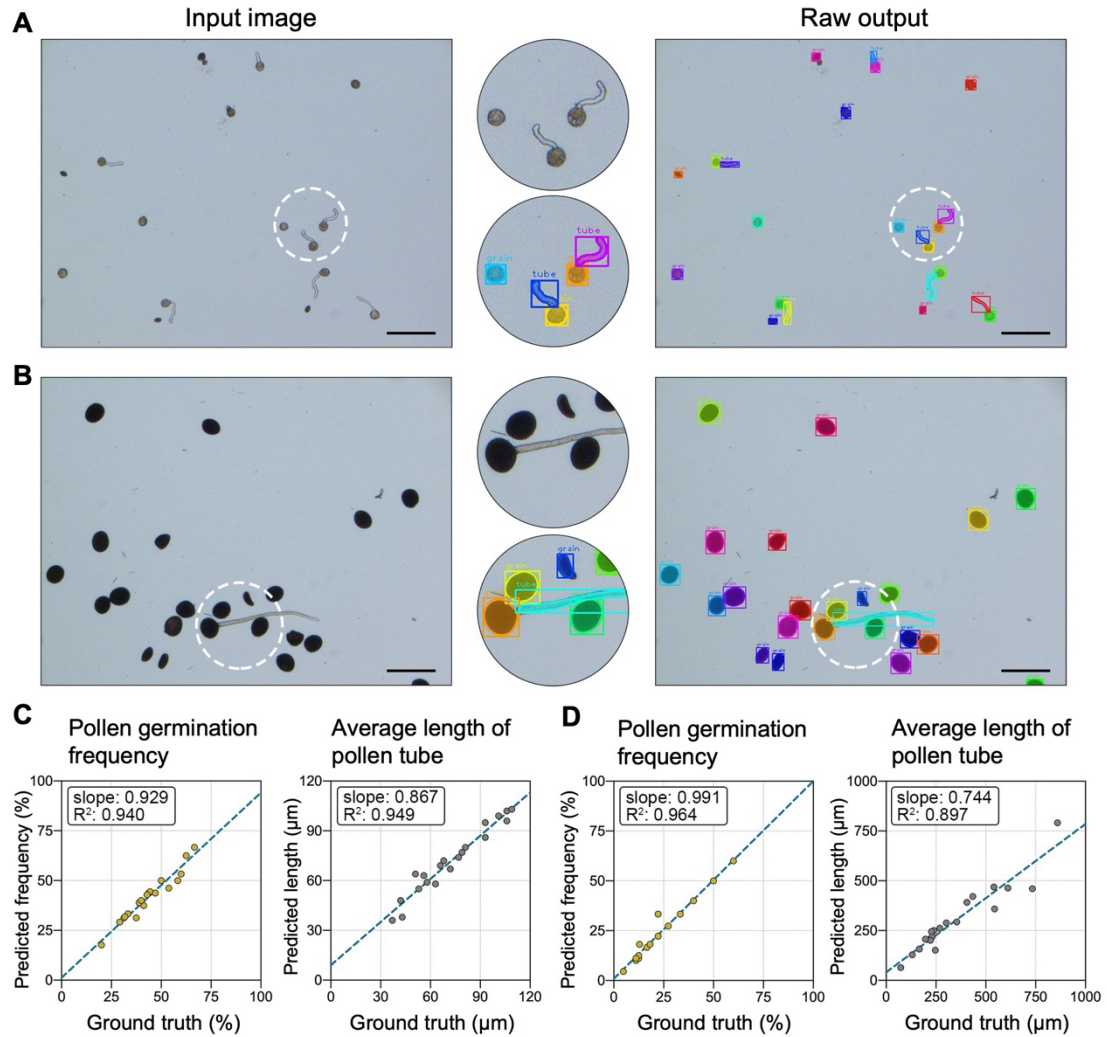


Fig. 2.10. Model generalization test of petunia and plantain lily pollen grains.

Input and raw output images of petunia (A) and plantain lily (B) pollen grains. Scatter plots of the results compared to the ground truth for pollen germination frequency and average pollen tube length of petunia (C) and plantain lily (D). Each dot represents an original microscopic image. The dark green dotted lines represent the linear regression equation. Scale bars = 200 μm .

This figure is from “Deep learning-based high-throughput detection of in vitro germination to assess pollen viability from microscopic images” (DOI: 10.1093/jxb/erad315) © 2023 by Mengwei Zhang, Jianxiang Zhao, Yoichiro Hoshino licensed under CC BY 4.0.

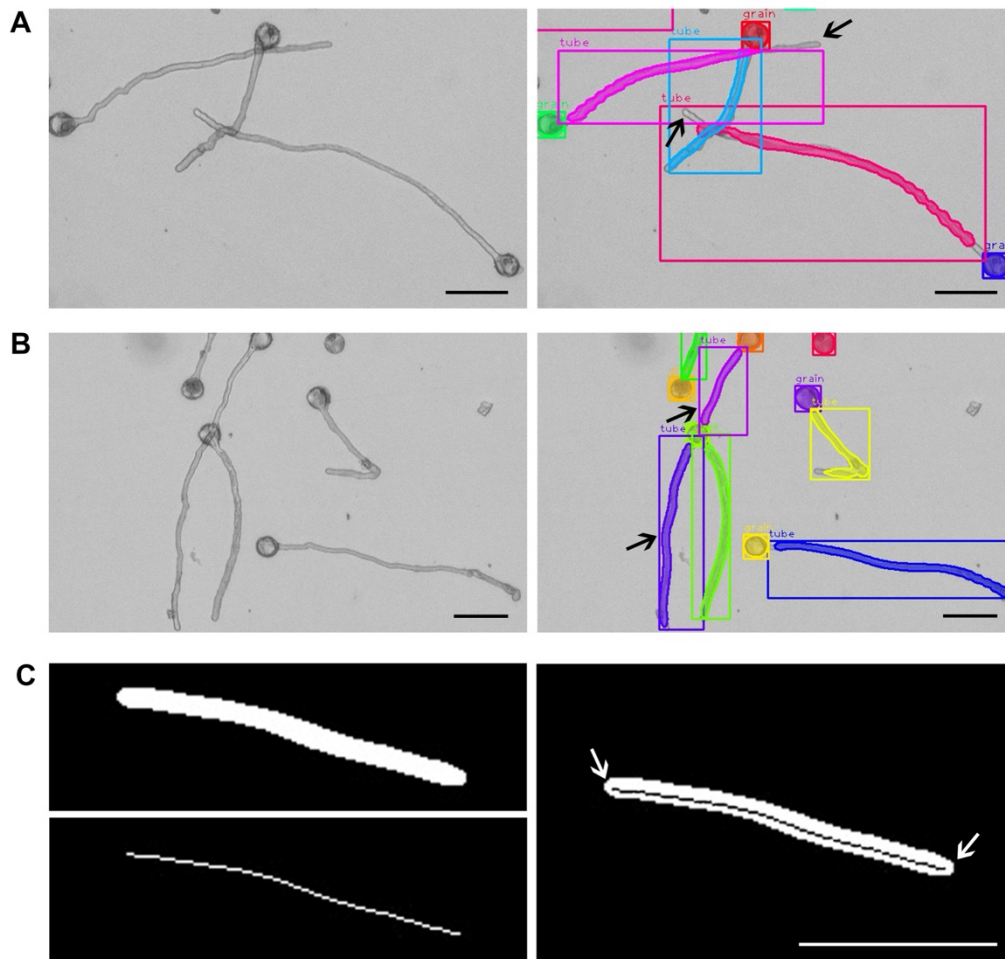


Fig. 2.11. Three types of errors that cause the detected pollen tube length to be smaller than the ground truth.

(A) Partial detection. (B) One pollen tube detected as two. (C) Mask and skeletonized line of a pollen tube (left side); the combination of mask and skeletonized line (right side).

Arrows indicate the locations of the error. Scale bars = 100 μm .

This figure is from “Deep learning-based high-throughput detection of in vitro germination to assess pollen viability from microscopic images” (DOI: 10.1093/jxb/erad315) © 2023 by Mengwei Zhang, Jianxiang Zhao, Yoichiro Hoshino licensed under CC BY 4.0.

Chapter 3: Morphology-based prediction of pollen germinability using deep-learning techniques in *Paeonia suffruticosa*

3.1 Introduction

The Mask R-CNN model discussed in Chapter 2 demonstrated robust generalizability, making it suitable for most pollen viability assessment needs. However, it has limitations. For example, when the pollen grain density is too high in *in vitro* culture experiments, the detection accuracy of the model decreases significantly due to the excessive intertwining of pollen tubes after germination.

Wodehouse (1935) refers the pollen morphological changes during dehydration, dispersal, and rehydration as “harmomegathy”. Pollen grains dehydrate gradually during development to enhance their adaptability to environmental changes (Pacini et al., 2006). The water content of pollen reaches a minimum at maturity, accompanied by a reduction in volume and morphological alterations (Nepi et al., 2001). This state enhances the resilience of pollen, ensuring a high success rate of long-distance pollination under natural conditions.

The morphology of pollen grains changes when they land on a compatible stigma or encounters a high-humidity environment due to hydration (Heslop-Harrison, 1979a; Heslop-Harrison, 1987). The shape of pollen varies among different plant species. For instance, tree peony pollen is dark and ellipsoidal when dehydrated, and transforms to light and spherical upon full hydration, because of the significant increase in the volume of mature and healthy pollen grains. In contrast, immature or defective pollen shows morphological differences from healthy mature pollen, remaining shriveled or incompletely hydrated (Langedijk et al., 2023).

The morphological differences between mature, viable pollen grains and immature, non-viable pollen grains upon hydration facilitates their classification using a computer vision system based on deep learning. The deep-learning model YOLO finds patterns and connections in a wide range of image data (Wang et al., 2023). In this chapter, I describe the development of a deep-learning model based on YOLOv7, which can predict pollen germination frequency based on the morphological characteristics of non-germinated hydrated pollen grains.

3.2 Materials and methods

3.2.1 Plant material and culture conditions

The tree peony plants utilized in this study were cultivated under field conditions in the Field Science Center for Northern Biosphere, Hokkaido University, Japan. Mature anthers from four cultivars ('Shimadaijin', 'Shimanishiki', 'Houki', and 'Ten-i') were collected prior to dehiscence. Subsequently, the anthers were dried at 25 °C for 12 h and stored in a –20 °C fridge until further use.

In vitro pollen culture was performed according to Zhang et al. (2023). The anthers were vortex-mixed with 1.5 mL of liquid medium in a 2-mL centrifuge tube. The resulting suspension of pollen grains was then transferred to a 3.5-cm plastic petri dish for microscopic observation. The composition of the liquid medium was identical to that described in Chapter 2. The culture medium was sterilized via autoclaving and was subsequently stored at 4 °C before use.

3.2.2 Morphological analysis of hydrated pollen grains

Preliminary experiments showed that during *in vitro* pollen culture, the tree peony pollen grains typically achieve full hydration approximately 60 s after immersion in the liquid medium (Fig. 3.1A). Images of hydrated pollen grains were captured using a light microscope (DS-Fi3; Nikon) with a 5× objective lens. The morphological parameters, including shape, size, and color, of hydrated pollen grains were measured using ImageJ 2.9.0 (Abràmoff et al., 2004). Pollen color was determined by calculating the mean gray value. Pollen shape was analyzed by fitting an ellipse to obtain the aspect ratio of the major and minor axes. Correlation analysis and principal component analysis (PCA) were then performed using these three parameters of 180 ‘Shimadaijin’ pollen grains on R studio.

3.2.3 Image acquisition and labeling of the deep-learning model

Given the morphological differences, such as variations in shape, size, and color, in tree peony pollen before and after full hydration, this study aimed to develop a novel pollen quality evaluation system based on deep learning to determine the germinability of pollen grains. I selected the YOLOv7 model for this task.

Image acquisition and labeling was performed as follows (Fig. 3.1B): during *in vitro* pollen culture, an initial image, M_0 , was captured once the pollen grains were fully hydrated. Subsequently, a follow-up image, M_1 , was captured at the same location once the pollen germination frequency had stabilized, aiming to collect information on pollen germination. This information was then utilized to label the original image, M_0 , using the local annotation tool LabelImg (<https://github.com/tzutalin/labelImg>). All the images were prepared at a fixed size of 2880×2048 pixels.

3.2.4 Determination of the optimal timing for capturing image M_I

To identify the optimal time for capturing image M_I , which contained the pollen germination information for labeling, I performed a pollen culture experiment on four tree peony cultivars to determine when each cultivar reaches stable pollen germination frequency. Images were captured at 0.5, 1, 2, and 3 h after the *in vitro* pollen culture. Pollen germination frequency was calculated in five random microscopic fields, and these experiments were repeated six times. The resulting data represent the mean of these experiments, with the standard error calculated.

3.2.5 Model training and evaluation

For each cultivar, 80 sets of images were captured using the acquisition approach (Fig. 3.1B); 40 were used for model training, and the remaining 40 were included in the test dataset. The YOLO model was trained over 800 epochs using 640 augmented training images. To determine the relationship between the predicted number and the actual number, linear regression was employed, followed by the calculation of the root mean squared error (RMSE).

Validation loss, mAP, recall, and precision were calculated to assess the model quality. The meanings of validation loss and mAP have been elucidated in Chapter 2. Recall refers to the proportion of actual positives that are correctly identified by the model (a), and precision refers to the proportion of positive predictions that are correct (b). A precision-recall curve is generated by plotting precision against recall at various thresholds (Everingham et al., 2010). Each precision-recall pair is represented as a point on the graph, and the curve is generated by connecting these points. The precision-recall curve is ideal for evaluating models in tasks that require high precision and recall.

$$Recall = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Negatives\ (FN)} \quad (a)$$

$$Precision = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Positives\ (FP)} \quad (b)$$

3.3 Results

3.3.1 Morphological analysis of pollen grains

Among the 180 randomly selected ‘Shimadaijin’ pollen grains, 108 germinated (with germinability), whereas 72 did not (without germinability). After hydration, the pollen grains exhibited considerable morphological variations. The morphological data analysis of these 180 hydrated pollen grains is presented in Fig. 3.2.

In the light microscopy images, the area distribution of hydrated pollen grains is widespread (Fig. 3.2A), ranging from 4 to 24 μm^2 , with grains below 12 μm^2 generally lacking germinability and those exceeding 14 μm^2 demonstrating germination potential. In Fig. 3.2B, the gray values of pollen are normally distributed. Apart from the interval where the gray values are below 60, pollen grains in all other intervals exhibit germination potential. Additionally, pollen grains with higher gray values tended to possess higher germinability. Most pollen grains were able to swell into a rounded shape by absorbing water, whereas only a small portion, which were inadequately hydrated, exhibited an elliptical or spindle shape (Fig. 3.2C). According to aspect ratio analysis, the subrounded shape is a typical stable feature among germinable pollen. Pollen grains that are fully hydrated do not necessarily possess germinability; however, all germinable pollen grains are fully hydrated. In contrast, inadequately hydrated pollen grains often fail to germinate.

The correlation analysis between pollen germinability and three morphological parameters is illustrated in Fig. 3.2D. Pollen germinability exhibited a strong positive

correlation with pollen grain area ($\rho = 0.81$), moderate positive correlation with gray value ($\rho = 0.40$), and moderate negative correlation with pollen grain aspect ratio ($\rho = -0.40$). In other words, pollen grains with larger areas, brighter colors, and nearly circular shapes under the optical microscope tended to possess higher germinability after hydration. Additionally, the pollen grain area showed a weak positive correlation with gray value ($\rho = 0.32$) and moderate negative correlation with aspect ratio ($\rho = -0.50$). Gray value also showed a weak negative correlation with aspect ratio ($\rho = -0.24$). These findings suggest that pollen grains with larger areas tended to exhibit brighter colors and circular shapes, whereas smaller pollen grains typically possessed darker colors and more elliptical shapes.

PCA revealed two distinctions in the morphology regarding pollen grains with or without germinability (Fig. 3.2E). Despite some overlap between the two groups, noticeable clustering was observed. A deep-learning model may be capable of classifying them based on these collective findings.

3.3.2 Determination of the optimal time point of image capturing

The pollen germination trend among the four tree peony cultivars followed a consistent pattern (Fig. 3.3), with all exhibiting rapid germination within the first 0.5 h and reaching stability by 1.5 h. This indicates that the images captured at 1.5 h accurately represent the pollen germination status. Therefore, these images were selected to provide the label information for the training dataset of the deep-learning model.

3.3.3 Model evaluation

After 800 epochs of training, the loss value and mAP tended to be stabilized (Fig. 3.4A). The model achieved the highest mAP of 0.979 and lowest validation loss of 0.057 on the validation dataset, with a 50% IoU threshold, demonstrating good performance.

The PR curves for both the classes showed the characteristics of high precision and recall (Fig. 3.4B). Furthermore, the area under the PR curve (AUC-PR) was high, indicating that the model performed well overall. These results demonstrated the capacity of the model to correctly identify a significant portion of instances while minimizing false positives.

3.3.4 Model application and evaluation

Pollen was classified into the following two categories: pollen grains with germinability and pollen grains without germinability. Fig. 3.5 provides an application example of the proposed model. Users simply need to prepare hydrated pollen microscopic images and input them into the model, which will then promptly output predictive results for the germinability of each pollen grain and perform counting. Based on these numbers, the predicted germination frequency can be calculated to evaluate pollen quality.

To validate the predictions, the output was compared against ground truth, specifically for the images of pollen captured at the optimal time point. The model exhibited a remarkable capacity to identify and segment each pollen grain, ensuring the precision of pollen grain counting. Furthermore, comparing the output with ground truth showed that the model effectively predicted pollen germination across the four cultivars of tree peony.

Precise counting of both pollen categories is essential for predicting the germination frequency. Fig. 3.6 presents the model evaluation results based on the test dataset. For the number of pollen grains with germinability (Fig. 3.6A) and without germinability (Fig. 3.6B), the R^2 for all four cultivars were close to 1 and RMSE were relatively low (Table 3.1). These indicate that the model was well-fitted and that there is a strong relationship between the predicted and actual values. In terms of pollen germination frequency (Fig. 3.6C), the lines connecting the actual and predicted values for most of the individual images approximate a horizontal alignment, indicating that the actual values were closely aligned with the predicted values. Furthermore, across repeated experiments involving 40 images per cultivar, no significant differences were observed between the actual and predicted values. This adequately demonstrated the stability of the model.

3.4 Discussion

In this study, I developed a deep-learning model, based on YOLOv7, which could predict the pollen germinability of tree peony pollen using microscopic images of hydrated pollen, with high accuracy.

In Fig. 3.1A, the tree peony pollen looks like a dark spindle before hydration, with clearly visible furrows on the pollen wall. Pollen can be categorized into two types, recalcitrant and orthodox, based on its desiccation tolerance (Bajaj, 1987). Pollen wall has aperture structures that allow pollen tube germination and material exchange with the external environment (Heslop-Harrison, 1979b). During harmomegathy, the pollen wall structure of orthodox pollen permits the formation of folds upon dehydration. These folds close the apertures inwardly, protecting against excessive water loss, which also enables the pollen grain to exhibit an elliptical shape with furrows to accommodate for the

decrease in cellular volume caused by water loss (Wodehouse, 1935; Katifori et al., 2010). After full hydration, the appearance of orthodox pollen becomes more rounded due to increased water content (Franchi et al., 2011). Fig. 3.1 shows the expanded pollen grains after hydration, with more distinct boundaries. The morphological differences between each pollen grain are more noticeable. Recalcitrant pollen lacks a furrow structure; thus, despite experiencing desiccation-induced shrinkage in volume, it does not exhibit significant changes in shape. Compared to recalcitrant pollen, the morphological differences of orthodox pollen before and after hydration are more pronounced (Pacini and Rudy, 2019). The pollen grains of tree peony are typical examples of orthodox pollen.

Mature and healthy tree peony pollen grains exhibit a dark olive-shaped morphology before hydration, transitioning to a larger, subrounded shape after hydration (Fig. 3.2), consistent with the identification results of the impedance flow cytometry based on pollen electrical properties (Langedijk et al., 2023). The mechanical stress generated during harmomegathy and maintenance of its full and rounded morphology need to be collectively supported by the complete pollen walls, plasma membrane, and protoplast (Blackmore and Barnes, 1986). Injured or dead pollen grains, due to structural damage, cannot fully absorb water and expand, leading to morphological differences compared to healthy pollen grains capable of germination.

The varying water content ultimately leads to morphological differences of hydrated tree peony pollen. Compared with live pollen, injured or dead pollen, due to membrane incompleteness, exhibits a significantly smaller volume after hydration (Ascari et al., 2020). The furrow structures unique to orthodox pollen remain folded due to incomplete hydration, resulting in an elliptical rather than a subrounded shape.

Furthermore, the variation in pollen grayscale value is attributed to differences in water content, leading to disparities in light refraction (Zhang et al., 2018).

In the PCA based on the three morphological features of hydrated pollen shown in Fig. 3.2E, some pollen grains that ultimately did not germinate were classified into the germinable group. Carrizo García et al. (2017) described the metabolic state of individual pollen grains as being asynchronous during anther opening. In certain instances, pollen may not germinate despite exhibiting the appearance of germinable pollen (Fig. 3.7A); this may be due to the asynchronous development of pollen cells. Other pollen grains that fail to germinate can be classified into the following two categories: mature but dead pollen (Fig. 3.7B) and aberrant pollen (Fig. 3.7C). Similar morphologies were observed by Iwahori (1965) in tomato pollen subjected to high-temperature stress. These two types of pollen do not undergo full hydration, remain wrinkled, and exhibit a dark color. They have distinct morphological differences, facilitating their identification by the proposed model.

Notably, this experimental approach employed a 5× objective lens for phenotype analysis and model training, potentially overlooking critical features, such as the surface texture of pollen grains. Moreover, since microscopic images are naturally flat, using 3D pollen imaging can provide a deeper understanding in future studies.

3.5 Conclusion

A YOLOv7 model was developed to predict pollen germinability based on the morphological characteristics of hydrated tree peony pollen grains. The novelty of this study lies in the fact that the proposed model can be used for the non-invasive inspection of pollen viability before germination. This approach significantly reduces the time

required for *in vitro* pollen culture, enabling the rapid and accurate prediction of pollen germination frequency. Importantly, it also preserves the viability of the tested pollen, ensuring the feasibility of subsequent experiments.

Table 3.1. Linear regression for four cultivars.

	Cultivar	Equation	R^2	RMSE
Pollen grains with germinability	‘Houki’	$y = 1.030x + 0.066$	0.993	1.360
	‘Shimanishiki’	$y = 0.990x + 1.715$	0.974	2.092
	‘Shimadaijin’	$y = 0.996x + 0.122$	0.849	2.264
	‘Ten-i’	$y = 0.964x + 0.922$	0.922	1.245
Pollen grains without germinability	‘Houki’	$y = 0.960x + 0.456$	0.990	1.449
	‘Shimanishiki’	$y = 0.961x - 0.800$	0.903	1.981
	‘Shimadaijin’	$y = 0.811x + 3.627$	0.855	2.179
	‘Ten-i’	$y = 1.003x - 0.569$	0.979	1.107

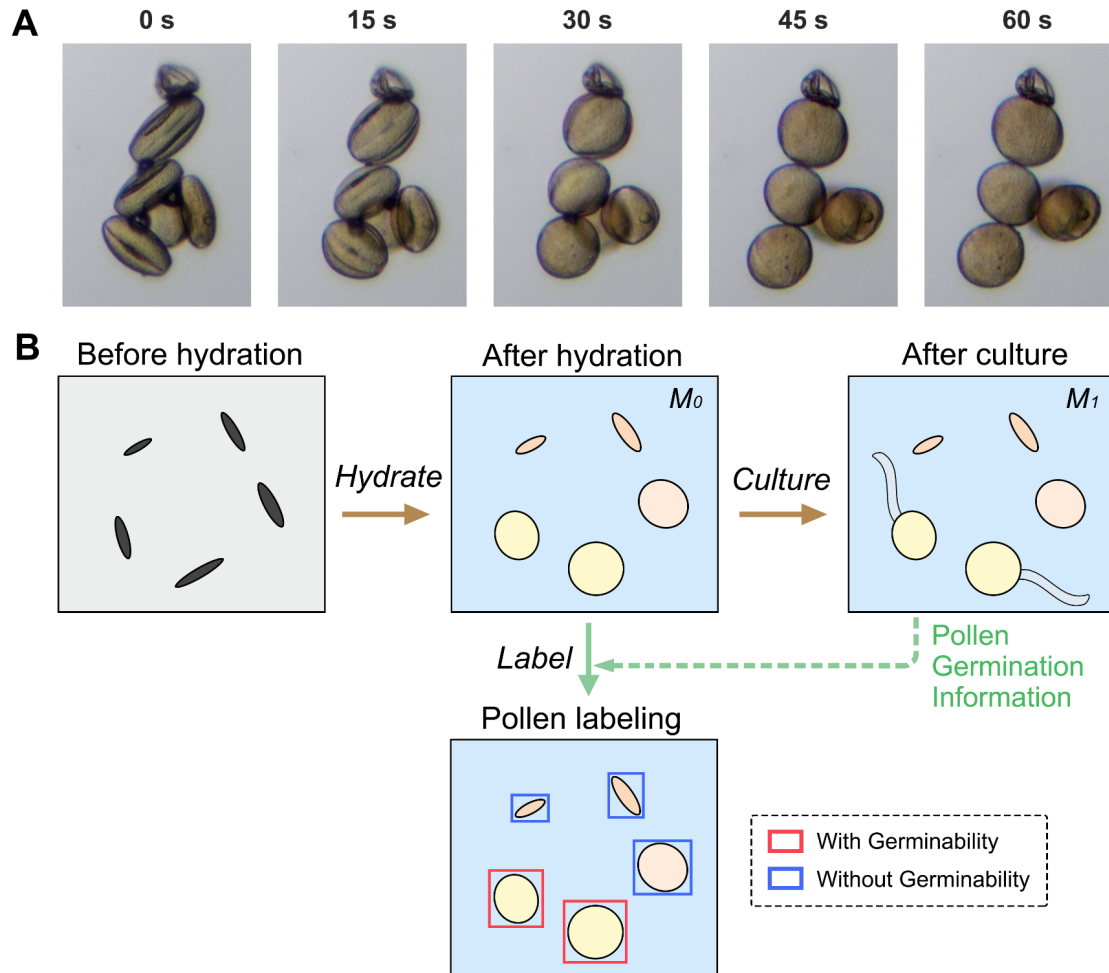


Fig. 3.1. Labeling of the deep-learning training dataset using the morphological differences of hydrated tree peony pollen.

(A) Hydration of tree peony pollen for 60 s. (B) Image acquisition and labeling for the YOLO model.

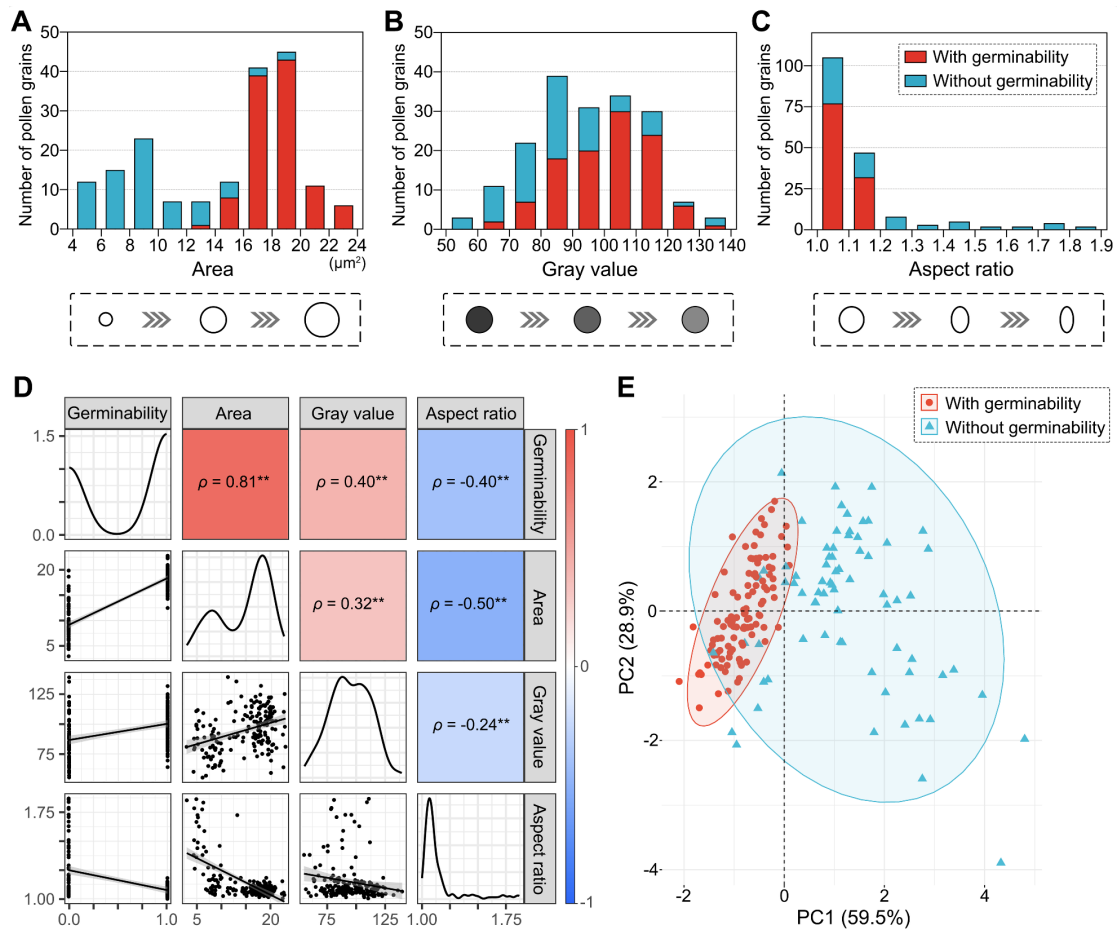


Fig. 3.2. Morphological data analysis of 180 hydrated 'Shimadaijin' pollen grains.

(A) Distribution chart of pollen area. (B) Distribution chart of pollen gray value. (C) Distribution chart of pollen aspect ratio. The dashed boxes below each chart show schematic diagrams of how the traits change along the x-axis. (D) Scatter plots, line graphs, and correlation coefficients (ρ : Spearman) for four variables (n = 180). (E) Factor map of the PCA performed using three variables of 180 'Shimadaijin' pollen grains. Two cluster groups were identified, corresponding to pollen grains with and without germinability. **, $P < 0.01$.

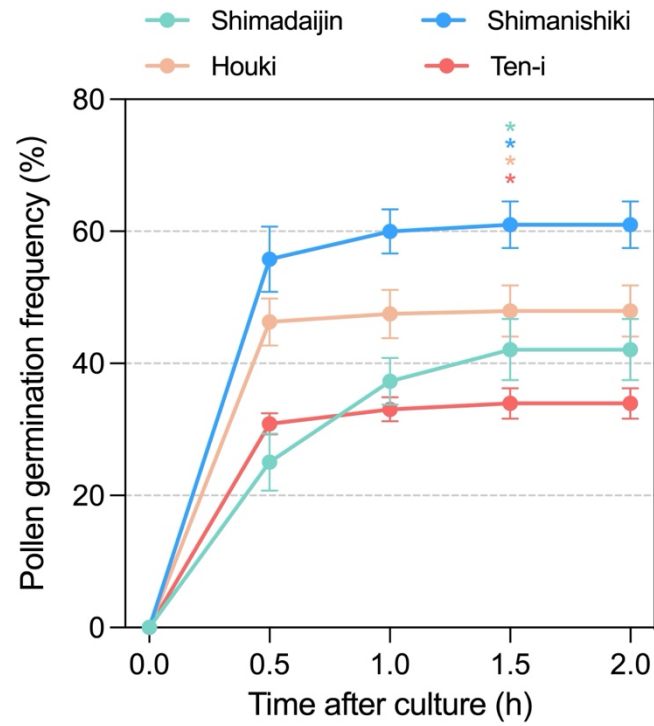


Fig. 3.3. Pollen germination frequency among four tree peony cultivars at 0.5, 1, 1.5, and 2 h after pollen culture.

The asterisks indicate the minimum time required for each cultivar to reach a stable germination frequency.

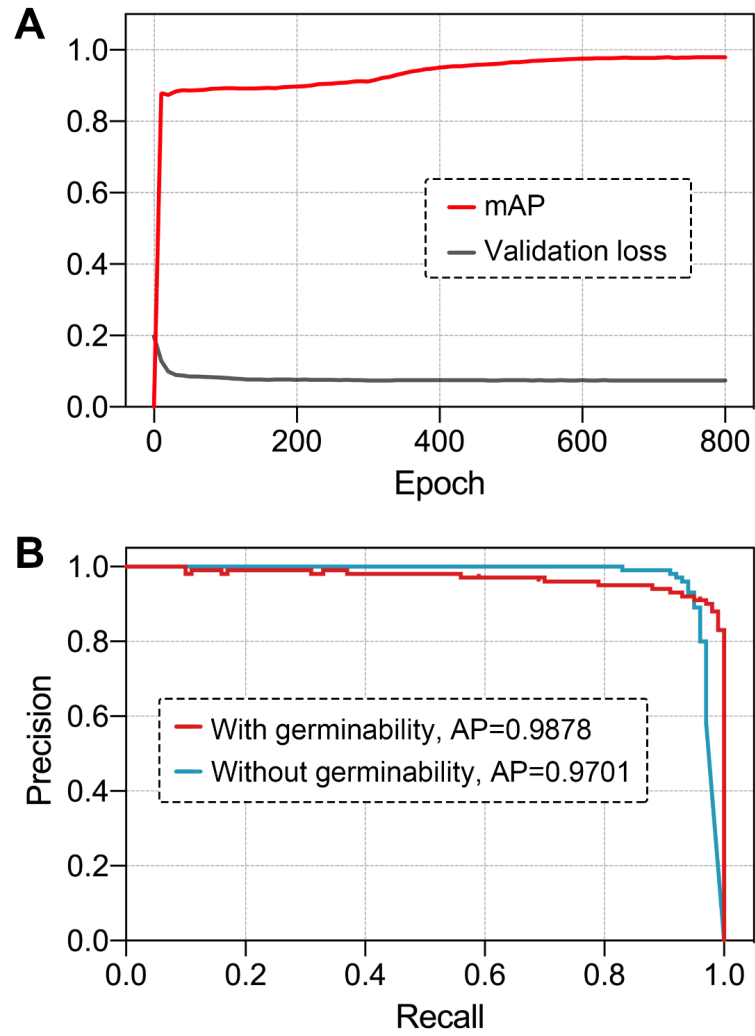


Fig. 3.4. Evaluation of the proposed YOLOv7 model using the validation dataset. (A) Validation loss and mAP at an IoU threshold of 50%. (B) Precision-recall curve under an IoU threshold of 50%.

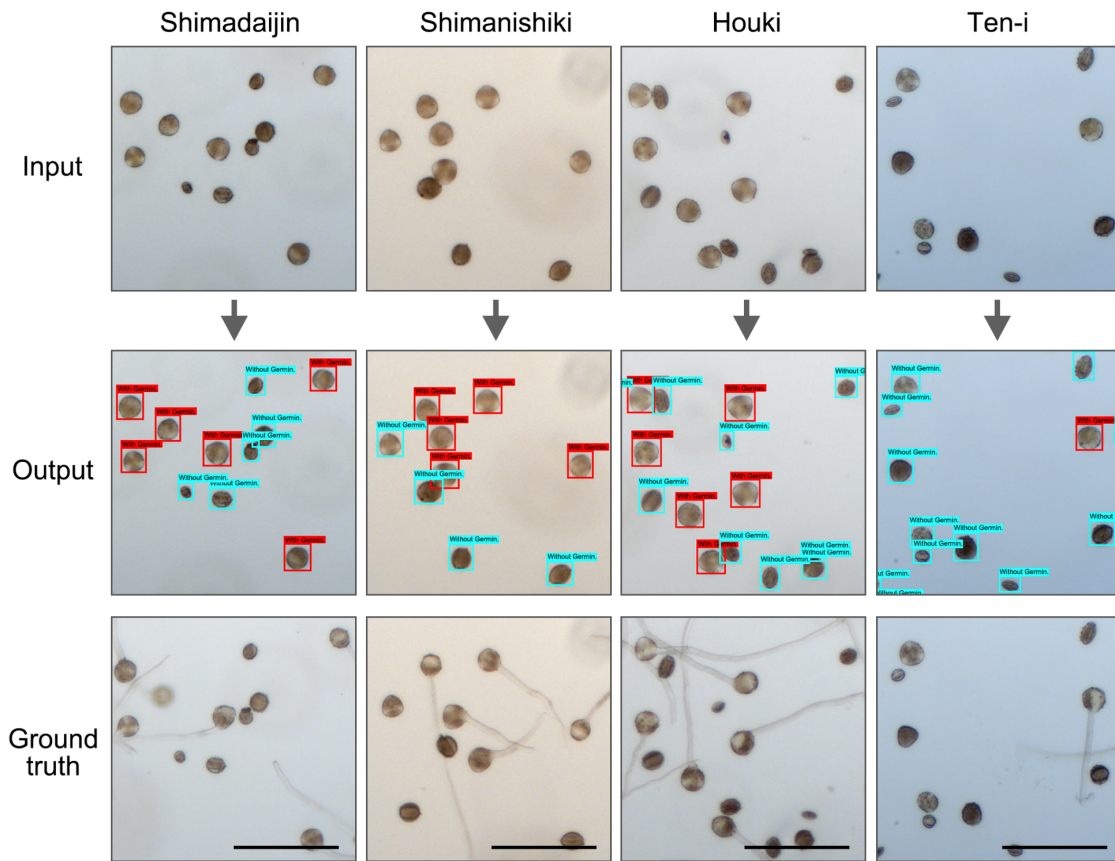


Fig. 3.5. YOLO model predictions and ground truth for four cultivars of tree peony. The red boxes indicate pollen grains with germinability, whereas the blue boxes indicate those without germinability. Scale bars = 1 mm.

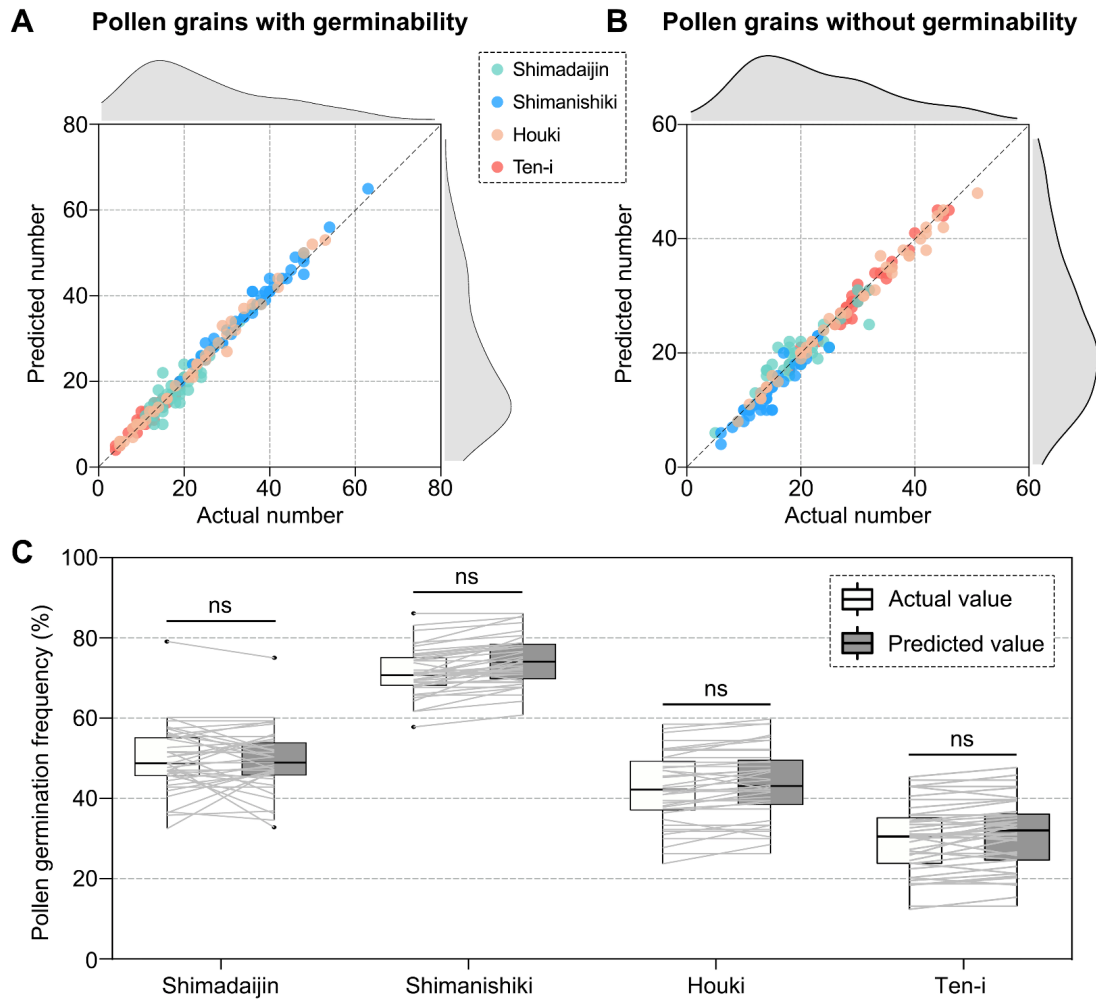


Fig. 3.6. Evaluation of the proposed model based on the test dataset.

Scatter plot of the model predicted number of pollen grains (A) with germinability and (B) without germinability against the actual values. The black dashed lines represent the function $y = x$. (C) Comparison between the actual and model-predicted germination frequency of four tree peony cultivars. The connection lines between the actual and predicted value link the same images. Significant differences were analyzed using a t -test; ns, not significant.



Fig. 3.7. Three types of pollen grains that cannot germinate.

(A) Not fully matured pollen. (B) Mature dead pollen. (C) Aberrant pollen.

Chapter 4: General discussion

4.1 Advantages of the deep-learning technique in a pollen viability assessment

This study proposed two deep-learning models based on Mask R-CNN and YOLOv7, to assess pollen viability by detecting the microscopic images of tree peony pollen.

As mentioned in Chapter 1, several approaches for assessing pollen viability require cytochemical staining, such as Alexander (Alexander, 1969), FDA (Heslop-Harrison and Heslop-Harrison, 1970), FDA-PI double (Jones and Senft, 1985), and H₂DCFDA staining (Luria et al., 2019; Rutley and Miller, 2020). Presently, multi-omics analyses are widely employed in plant science (Yang et al., 2021). These analyses demand a very high level of purity in the experimental materials, and using stains negatively impacts the subsequent analysis of the tested pollen. In contrast, the deep learning-based image detection method is a non-invasive method. Additionally, the liquid culture medium developed by Hirano and Hoshino (2009, 2010) was used in this study, facilitating the subsequent utilization of the detected pollen.

4.2 Novelty of the approaches

The Mask R-CNN and YOLO models belong to different model types. Mask R-CNN is an instance segmentation model (He et al., 2017), whereas YOLO is an object detection model (Redmon et al., 2016), leading to logical differences in the experimental workflow.

The Mask R-CNN model proposed in Chapter 2 can calculate the pollen germination frequency by separately counting pollen grains and tubes. Additionally, through code modification, the model can also use the collected pollen tube masks to

estimate the average pollen tube length. Therefore, the procedure detailed in this chapter is the first to use the Mask R-CNN model for evaluating pollen viability.

The YOLOv7 model proposed in Chapter 3 can determine the germination potential of hydrated pollen grains based on their morphology. In a related study, Yamazaki et al. (2023) used the YOLOv5 model to calculate pollen germination frequency. However, this model could only recognize pollen after *in vitro* culture. In contrast, the proposed YOLOv7 model could make accurate and rapid predictions without the need for prior *in vitro* pollen culture.

4.3 Limitations and prospects

The execution time of the Mask R-CNN and YOLO models vary significantly due to their varying architectures (Soviany and Ionescu, 2018). The proposed Mask R-CNN represents a two-stage detector based on the R-CNN algorithm incorporating a region proposal network (RPN). Furthermore, the post-processing steps, such as “pollen tube mask extraction” and “mask skeletonization,” contribute to a considerable computational overhead, resulting in an approximate execution time of 30 s per image. Although notably faster than manual counting methods, the execution time remains significantly longer compared to the one-stage detector, YOLOv7, which takes only approximately 0.1 s per image.

The YOLO model is renowned for its fast execution time, making it a preferred choice for real-time detection tasks, such as in-field weed recognition (Dang et al., 2023) or assisted fruit harvesting (Wang et al., 2022). Future studies can couple the proposed YOLOv7 model with appropriate hardware to achieve real-time detection, thereby accomplishing the quality assessment of freshly collected pollen samples in the shortest

possible time. Notably, the proposed Mask R-CNN model was designed to detect pollen grains and tubes, which gives it a wide range of applicability. This model can be applied to the pollen of various plants through transfer learning (Tan et al., 2018). However, the proposed YOLOv7 model was specifically tailored for the detection of tree peony pollen, and therefore, has limited generalizability. This model needs to be retrained to be used in other plant species. Despite these limitations, using YOLOv7 in this study demonstrated significant advancements in non-invasive pollen viability assessments, offering a valuable tool in plant sciences, particularly in applications directly related to the specific characteristics of tree peony pollen.

The two deep learning-based systems proposed in this study offer a non-invasive, accurate, and rapid approach to detect pollen viability. This technical support can be highly beneficial for breeding tree peonies and potentially other plant species.

References

- Abdul-Baki AA. 1992. Determination of pollen viability in tomatoes. *Journal of the American Society for Horticultural Science* 117, 473-476.
- Abràmoff MD, Magalhães PJ, Ram SJ. 2004. Image processing with imageJ. *Biophotonics International* 11, 36-41.
- Albawi S, Mohammed TA, Al-Zawi S. 2017. Understanding of a convolutional neural network. In: *Proceedings of 2017 International Conference on Engineering and Technology*. Antalya, Turkey, Aug 21-23. 1-6.
- Alexander MP. 1969. Differential staining of aborted and nonaborted pollen. *Stain Technology* 44, 117-122.
- Ascari L, Novara C, Dusio V, Oddi L, Siniscalco C. 2020. Quantitative methods in microscopy to assess pollen viability in different plant taxa. *Plant Reproduction* 33, 205-219.
- Bajaj YPS. 1987. Cryopreservation of pollen and pollen embryos, and the establishment of pollen banks. *International Review of Cytology* 107, 397-420.
- Blackmore S, Barnes SH. 1986. Harmomegathic mechanisms in pollen grains. In: Blackmore S, Ferguson IK. eds. *Pollen and spores: forms and function*. London: Academic Press, 137-149.
- Boavida LC, McCormick S. 2007. Temperature as a determinant factor for increased and reproducible *in vitro* pollen germination in *Arabidopsis thaliana*. *Plant Journal* 52, 570-582.
- Brunet J, Ziobro R, Osvatic J, Clayton MK. 2019. The effects of time, temperature and plant variety on pollen viability and its implications for gene flow risk. *Plant Biology* 21, 715-722.
- Burke IC, Wilcut JW, Allen NS. 2007. Viability and *in vitro* germination of Johnsongrass (*Sorghum Halepense*) pollen. *Weed Technology* 21, 23-29.

- Carrizo García C, Nepi M, Pacini E. 2017. It is a matter of timing: asynchrony during pollen development and its consequences on pollen performance in angiosperms—a review. *Protoplasma* 254, 57-73.
- Cheng F. 2007. Advances in the breeding of tree peonies and a cultivar system for the cultivar group. *International Journal of Plant Breeding* 1, 89-104.
- Costa L, Ampatzidis Y, Rohla C, Maness N, Cheary B, Zhang L. 2021. Measuring pecan nut growth utilizing machine vision and deep learning for the better understanding of the fruit growth curve. *Computers and Electronics in Agriculture* 181, 105964.
- Covey PA, Subbaiah CC, Parsons RL, Pearce G, Lay FT, Anderson MA, Ryan CA, Bedinger PA. 2010. A pollen-specific RALF from tomato that regulates pollen tube elongation. *Plant Physiology* 153, 703-715.
- Crow JF. 1994. Advantages of sexual reproduction. *Developmental Genetics* 15, 205-213.
- Dafni A, Firmage D. 2000. Pollen viability and longevity: Practical, ecological and evolutionary implications. *Plant Systematics and Evolution* 222, 113-132.
- Dang F, Chen D, Lu Y, Li Z. (2023). YOLOWeeds: a novel benchmark of YOLO object detectors for multi-class weed detection in cotton production systems. *Computers and Electronics in Agriculture* 205: 107655.
- Dang LM, Wang H, Li Y, Park Y, Oh C, Nguyen TN, Moon H. 2022. Automatic tunnel lining crack evaluation and measurement using deep learning. *Tunnelling and Underground Space Technology* 124: 104472.
- Dunker S, Motivans E, Rakosy D, Boho D, Mäder P, Hornick T, Knight TM. 2021. Pollen analysis using multispectral imaging flow cytometry and deep learning. *New Phytologist* 229, 593-606.
- Dupl'áková N, Dobrev PI, Reňák D, Honys D. 2016. Rapid separation of *Arabidopsis* male gametophyte developmental stages using a Percoll gradient. *Nature Protocols* 11, 1817-1832.
- Dutta SK, Srivastav M, Chaudhary R, Lal K, Patil P, Singh SK, Singh AK. 2013. Low temperature storage of mango (*Mangifera indica* L.) pollen. *Scientia Horticulturae* 161, 193-197.

- Esser K. 1955. Pollen tubes in the style after chloral hydrate and aceto-carmin. *Stain Technology* 30, 195-197.
- Everingham M, Van Gool L, Williams CK, Winn J, Zisserman A. 2010. The pascal visual object classes (voc) challenge. *International Journal of Computer Vision* 88, 303-338.
- Fahad S, Hussain S, Saud S, Khan F, Hassan S, Amanullah, Nasim W, Arif M, Wang F, Huang J. 2016. Exogenously applied plant growth regulators affect heat-stressed rice pollens. *Journal of Agronomy and Crop Science* 202, 139-150.
- Franchi GG, Piotto B, Nepi M, Baskin CC, Baskin JM, Pacini E. 2011. Pollen and seed desiccation tolerance in relation to degree of developmental arrest, dispersal, and survival. *Journal of Experimental Botany*, 62, 5267-5281.
- Fritz SE, Lukaszewski AJ. 1989. Pollen longevity in wheat, rye and triticale. *Plant Breeding* 102, 31-34.
- García-Fortea E, García-Pérez A, Gimeno-Páez E, Sánchez-Gimeno A, Vilanova S, Prohens J, Pastor-Calle D. 2020. A deep learning-based system (Microscan) for the identification of pollen development stages and its application to obtaining doubled haploid lines in eggplant. *Biology* 9: 272.
- Go RC, Daga RRM, Gerona MEB. 2019. Image analysis protocol for pollen viability identification on selected genotypes of rice (*Oryza sativa* L.). In: 2019 IEEE 4th International Conference on Computer and Communication Systems. Singapore, Feb 23-25. 248-252.
- He K, Gkioxari G, Dollár P, Girshick R. 2017. Mask R-CNN. In: Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy, Oct 22-29. 2980-2988.
- Heslop-Harrison J, Heslop-Harrison Y. 1970. Evaluation of pollen viability by enzymatically induced fluorescence; intracellular hydrolysis of fluorescein diacetate. *Stain Technology* 45, 115-120.
- Heslop-Harrison J. 1979a. An interpretation of the hydrodynamics of pollen. *American Journal of Botany* 66, 737-743.

- Heslop-Harrison J. 1979b. Pollen walls as adaptive systems. *Annals of the Missouri Botanical Garden* 66, 813-829.
- Helpson-Harrison J, Heslop-Harrison Y, Shivanna KR. 1984. The evaluation of pollen quality, and a further appraisal of the fluorochromatic (FCR) test procedure. *Theoretical and Applied Genetics* 67, 367-375.
- Heslop-Harrison J. 1987. Pollen germination and pollen-tube growth. *International Review of Cytology* 107, 1-78.
- Heidmann I, Schade-Kampmann G, Lambalk J, Ottiger M, Berardino MD. 2016. Impedance flow cytometry: A novel technique in pollen analysis. *PLoS ONE* 11, 1-15.
- Hirano T, Hoshino Y. 2009. Detection of changes in the nuclear phase and evaluation of male germ units by flow cytometry during *in vitro* pollen tube growth in *Alstroemeria aurea*. *Journal of Plant Research* 122, 225-234.
- Hirano T, Hoshino Y. 2010. Sperm dimorphism in terms of nuclear shape and microtubule accumulation in *Cyrtanthus mackenii*. *Sexual Plant Reproduction* 23, 153-162.
- Hoshino Y, Eiraku N, Ohata Y, Komai F. 2016. Dynamics of nuclear phase changes during pollen tube growth by using *in vitro* culture in *Petunia*. *Scientia Horticulturae* 210, 143-149.
- Impe D, Reitz J, Köpnick C, Rolletschek H, Börner A, Senula A, Nagel M. 2020. Assessment of pollen viability for wheat. *Frontiers in Plant Science* 10, 1-13.
- Iwahori S. 1965. High temperature injuries in tomato. IV. Development of normal flower buds and morphological abnormalities of flower buds treated with high temperature. *Journal of the Japanese Society for Horticultural Science* 34, 33-41.
- Jiang Y, Lahlali R, Karunakaran C, Kumar S, Davis AR, Bueckert RA. 2015. Seed set, pollen morphology and pollen surface composition response to heat stress in field pea. *Plant, Cell and Environment* 38, 2387-2397.

- Jones KH, Senft JA. 1985. An improved method to determine cell viability by simultaneous staining with fluorescein diacetate-propidium iodide. *Journal of Histochemistry and Cytochemistry* 33, 77-79.
- Katifori E, Alben S, Cerda E, Nelson DR, Dumais J. 2010. Foldable structures and the natural design of pollen grains. *Proceedings of the National Academy of Sciences* 107, 7635-7639.
- Kaya A, Keceli AS, Catal C, Yalic HY, Temucin H, Tekinerdogan B. 2019. Analysis of transfer learning for deep neural network-based plant classification models. *Computers and Electronics in Agriculture* 158, 20-29.
- Khatun S, Flowers TJ. 1995. Effects of salinity on seed set in rice. *Plant, Cell and Environment* 18, 61-67.
- Kubera E, Kubik-Komar A, Kurasiński P, Piotrowska-Weryszko K, Skrzypiec M. 2022. Detection and recognition of pollen grains in multilabel microscopic images. *Sensors* 2222: 2690.
- Langedijk NS, Kaufmann S, Vos E, Ottiger T. 2023. Evaluation of methods to assess the quality of cryopreserved Solanaceae pollen. *Scientific Reports* 13: 7344.
- Lim EC, Kim J, Park J, et al. 2020. DeepTetrad: high-throughput image analysis of meiotic tetrads by deep learning in *Arabidopsis thaliana*. *Plant Journal* 101, 473-483.
- Luria G, Rutley N, Lazar I, Harper JF, Miller G. 2019. Direct analysis of pollen fitness by flow cytometry: implications for pollen response to stress. *Plant Journal* 98, 942-952.
- Marie R, Labbani-Igbida O, Mouaddib EM. 2016. The Delta Medial Axis: A fast and robust algorithm for filtered skeleton extraction. *Pattern Recognition* 56, 26-39.
- Mazzeo A, Palasciano M, Gallotta A, Camposeo S, Pacifico A, Ferrara G. 2014. Amount and quality of pollen grains in four olive (*Olea europaea* L.) cultivars as affected by 'on' and 'off' years. *Scientia Horticulturae* 170, 89-93.
- Nepi M, Franchi GG, Padni E. 2001. Pollen hydration status at dispersal: cytophysiological features and strategies. *Protoplasma* 216, 171-180.

- Neyshabur B, Sedghi H, Zhang C. 2020. What is being transferred in transfer learning? In: Proceedings of the 34th Conference on Neural Information Processing Systems. Vancouver, Canada, Dec 6-12. 512-523.
- Okuyama Y, Kato M. 2009. Unveiling cryptic species diversity of flowering plants: successful biological species identification of Asian *Mitella* using nuclear ribosomal DNA sequences. BMC Evolutionary Biology 9, 1-16.
- Pacini E, Franchi GG, Lisci M, Nepi M. 1997. Pollen viability related to type of pollination in six angiosperm species. Annals of Botany 80, 83-87.
- Pacini E, Guarnieri M, Nepi M. 2006. Pollen carbohydrates and water content during development, presentation, and dispersal: a short review. Protoplasma 228, 73-77.
- Pacini E, Rudy D. 2019. Pollen developmental arrest: maintaining pollen fertility in a world with a changing climate. Frontiers in Plant Science 10: 429534.
- Poidevin L, Forment J, Unal D, Ferrando A. 2021. Transcriptome and translome changes in germinated pollen under heat stress uncover roles of transporter genes involved in pollen tube growth. Plant, Cell and Environment 44, 2167-2184.
- Rang ZW, Jagadish SVK, Zhou QM, Craufurd PQ, Heuer S. 2011. Effect of high temperature and water stress on pollen germination and spikelet fertility in rice. Environmental and Experimental Botany 70, 58-65.
- Redmon J, Divvala S, Girshick R, Farhadi A. 2016. You only look once: Unified, real-time object detection. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, June 27-30. 779-788.
- Reichler SA, Torres J, Rivera AL, Cintolessi VA, Clark G, Roux SJ. 2009. Intersection of two signalling pathways: Extracellular nucleotides regulate pollen germination and pollen tube growth via nitric oxide. Journal of Experimental Botany 60, 2129-2138.
- Romero IC, Kong S, Fowlkes CC, Jaramillo C, Urban MA, Oboh-Ikuenobe F, D'Apolito C, Punyasena SW. 2020. Improving the taxonomy of fossil pollen using convolutional neural networks and superresolution microscopy. Proceedings of the National Academy of Sciences of the United States of America 117, 28496–28505.
- Roze D. 2012. Disentangling the benefits of sex. PLoS Biology 10: e1001321.

- Rutley N, Miller G. 2020. Large-scale analysis of pollen viability and oxidative level using H₂DCFDA-staining coupled with flow cytometry. In: *Methods in Molecular Biology* 2160, 167–79.
- Saeed K, Tabędzki M, Rybnik M, Adamski M. 2010. K3M: A universal algorithm for image skeletonization and a review of thinning techniques. *International Journal of Applied Mathematics and Computer Science* 20, 317-335.
- Schoch CL, Ciufo S, Domrachev M, Hotton CL, Kannan S, Khovanskaya R, Leipe D, Mcveigh R, O'Neill K, Robbertse B, Sharma S, Soussov V, Sullivan JP, Sun L, Turner S, Karsch-Mizrachi I. 2020. NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database* 2020: baaa062.
- Shivanna KR, Linskens HF, Cresti M. 1991. Pollen viability and pollen vigor. *Theoretical and Applied Genetics* 81, 38-42.
- Soviany P, Ionescu RT. 2018. Optimizing the trade-off between single-stage and two-stage deep object detectors using image difficulty prediction. In: *20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*, Timisoara, Romania, Sept. 20-23. 209-214.
- Stone JL, Thomson JD, Dent-Acosta SJ. 1995. Assessment of pollen viability in hand-pollination experiments: a review. *American Journal of Botany* 82, 1186-1197.
- Tan C, Sun F, Kong T, Zhang W, Yang C, Liu C. 2018. A survey on deep transfer learning. In: *Artificial Neural Networks and Machine Learning-ICANN 2018: 27th International Conference on Artificial Neural Networks*. Rhodes, Greece, Oct. 4-7. 270-279.
- Tan Z, Yang J, Li Q, Su F, Yang T, Wang W, Aierxi A, Zhang X, Yang W, Kong J, Min L. 2022. PollenDetect: An open-source pollen viability status recognition system based on deep learning neural networks. *International Journal of Molecular Sciences* 23: 13469.
- Tian Y, Yang G, Wang Z, Li E, Liang Z. 2020. Instance segmentation of apple flowers using the improved mask R-CNN model. *Biosystems Engineering* 193, 264-278.

- Toda Y, Okura F, Ito J, Okada S, Kinoshita T, Tsuji H, Saisho D. 2020. Training instance segmentation neural network with synthetic datasets for crop seed phenotyping. *Communications Biology* 3: 173.
- Trognitz BR. 1991. Comparison of different pollen viability assays to evaluate pollen fertility of potato dihaploids. *Euphytica* 56, 143-148.
- Uchida A, Abe A, Hoshino Y, Kunitake H. 2012. Liquid culture system for mature pollen in Hyuganatsu (*Citrus tamurana* hort. ex Tanaka). *Horticultural Research (Japan)* 11, 173-179.
- Wang Q, Lu L, Wu X, Li Y, Lin J. 2003. Boron influences pollen germination and pollen tube growth in *Picea meyeri*. *Tree Physiology* 23, 345-351.
- Wang CY, Bochkovski A, Liao HYM. 2023. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver, Canada, June 17-24. 7464-7475.
- Wang ZY, Ge Y, Scott M, Spangenberg G. 2004. Viability and longevity of pollen from transgenic and nontransgenic tall fescue (*Festuca arundinacea*) (Poaceae) plants. *American Journal of Botany* 91, 523-530.
- Wang Z, Jin L, Wang S, Xu H. 2022. Apple stem/calyx real-time recognition using YOLO-v5 algorithm for fruit automatic loading system. *Postharvest Biology and Technology* 185: 111808.
- Warman C, Sullivan CM, Preece J, Buchanan ME, Vejlupkova Z, Jaiswal P, Fowler JE. 2021. A cost-effective maize ear phenotyping platform enables rapid categorization and quantification of kernels. *Plant Journal* 106, 566-579.
- Wodehouse RP. 1935. *Pollen grains: Their structure, identification and significance in science and medicine*. New York, NY, USA: McGraw-Hill Co., Inc.
- Xie J, Fernandes SB, Mayfield-Jones D, Erice G, Choi M, Lipka AE, Leakey ADB. 2021. Optical topometry and machine learning to rapidly phenotype stomatal patterning traits for maize QTL mapping. *Plant Physiology* 187, 1462-1480.

- Yamazaki A, Takezawa A, Nagasaka K, Motoki K, Nishimura K, Nakano R. 2023. A simple method for measuring pollen germination rate using machine learning. *Plant Reproduction* 36, 355-364.
- Yang X, Li H, Yu Y, Luo X, Huang T, Yang X. 2018. Automatic pixel-level crack detection and measurement using fully convolutional network. *Computer-Aided Civil and Infrastructure Engineering* 33, 1090-1109.
- Yang Y, Saand MA, Huang L, Abdelaal WB, Zhang J, Wu Y, Li J, Sirohi MH, Wang F. 2021. Applications of multi-omics technologies for crop improvement. *Frontiers in Plant Science* 12: 563953.
- Zhang M, Zhao J, Hoshino Y. 2023. Deep learning-based high-throughput detection of *in vitro* germination to assess pollen viability from microscopic images. *Journal of Experimental Botany* 74, 6551-6562.
- Zhang S, Kim N, Yokoyama W, Kim Y. 2018. Effects of moisture content on mechanical properties, transparency, and thermal stability of yuba film. *Food Chemistry* 243, 202-207.

Summary

Paeonia suffruticosa, commonly known as tree peony, is a popular ornamental shrub. Various tree peony cultivars have been produced through cross-pollination and artificial selection. Breeding new tree peony cultivars typically involves hybridization and selecting seedlings resulting from cross-pollination. A critical factor in successful hybrid breeding is pollen viability, which determines the ability of pollen grains to fertilize the ovules effectively. Similar to other plants, high pollen viability correlates with successful seed production in tree peony.

Pollen viability can be assessed using various methods, each with its own limitations. For instance, staining methods damage the integrity of the pollen cell, hampering its use subsequent analysis. IFC requires specialized and costly equipment. *In vitro* pollen culture is useful because it facilitates evaluating pollen viability while retaining the germination ability. However, *in vitro* culture combined with manual counting of pollen tubes is extremely time-consuming. To address the above issues, this study proposed two deep learning-based systems as non-invasive alternatives.

The first system is developed from the Mask R-CNN model. The proposed model is proficient at instance segmentation of pollen grains and tubes, allowing for the precise detection and counting of individual pollen grains and tubes in microscopic images. By adding a synthetic dataset to the training dataset, the model can handle more complex scenarios of pollen tube crossing. Additionally, including post-processing codes allows the model to skeletonize pollen tube masks, enabling it to calculate pollen germination frequency and the average length of pollen tubes. This enhancement provides deeper insights into pollen germinability. At an IoU threshold of 50%, the model achieved a mAP of 0.949. For higher IoU thresholds, the model still performed commendably, with mAP

values of 0.820 for mAP_{70} and 0.633 for $\text{mAP}@ [0.5:0.95]$. In terms of accuracy, the model achieved 0.997 for pollen grains and 0.903 for pollen tubes at an IoU threshold of 50%. These metrics highlight the model's effectiveness in accurately segmenting pollen grains and tubes, confirming its reliability and robustness during pollen analysis. The generalizability of the proposed Mask R-CNN model extends beyond tree peony pollen. Its robust segmentation capabilities facilitate its application to other plant species. Furthermore, transfer learning can enhance this adaptability, where a model trained on tree peony pollen can be fine-tuned for pollen from other plant species with evidently different appearances. This feature highlights the model's potential to serve as a valuable tool in diverse research, enabling the assessment of pollen viability without requiring extensive reconfiguration.

The second system is based on You only look once (YOLO) v7, which is an object detection model known for its efficiency and speed. This model was used to predict the pollen germination potential based on their hydrated morphological characteristics. By capturing the images of fully hydrated pollen grains, the study analyzed key morphological features, such as size, color, and shape, to determine the viability of the pollen. Under the microscope, tree peony pollen grains with an area $>14 \mu\text{m}^2$ typically exhibit germinability, whereas those with an area $<12 \mu\text{m}^2$ do not. Correlation analysis indicated a strong positive relationship between the pollen area and their germinability ($\rho = 0.81$). Additionally, pollen grains with lighter colors (higher grayscale values) often possess higher germinability, with a moderate correlation coefficient of $\rho = 0.40$. The aspect ratio, which is related to shape (pollen grains closer to a circular shape have a lower aspect ratio), exhibited a moderate negative correlation with germinability ($\rho = -0.40$). Pollen grains that expanded into a circular shape upon hydration typically

demonstrated higher germinability. Overall, these findings suggest that the hydrated pollen grains exhibit distinct features that correlate with their germinability. Pollen grains with larger areas, brighter colors, and more circular shapes are more likely to germinate upon hydration, highlighting the significance of these morphological characteristics as indicators of pollen viability. Based on these characteristics, the proposed model effectively categorized pollen grains into the following two categories: pollen grains with germinability and those without germinability. The model achieved a high mAP of 0.979 and low validation loss of 0.057 at an IoU threshold of 50% after training for 800 epochs. Additionally, the precision-recall curves for both classes exhibited high precision and recall, with a notably high AUC-PR, indicating the model's effectiveness in accurately identifying instances while minimizing false positives. These results support strong model performance. However, the model's specificity to the dataset in this study suggests that further training and adjustments will be necessary to apply it broadly across different plant species. YOLOv7's speed facilitates real-time processing, making it suitable for time-sensitive experiments.

Developing deep learning-based systems to assess pollen viability is a promising advancement in plant research. These technological advancements enable the rapid assessment of the viability of the pollen grains of tree peony or other plant species, facilitating research on crop production and effective breeding. The ability to accurately assess pollen viability using deep-learning techniques offers significant advantages, including greater accuracy, reduced time, and labor, enabling subsequent research on the tested pollen, and potential for real-time analysis, over traditional methods. These achievements are particularly valuable in the fundamental research on plant reproduction. Elucidating the function of pollen in plant reproduction may be applied to stable crop

production and utilization of unused plant genetic resources for breeding. Integrating deep learning into this field may provide opportunities for enhancing breeding strategies and potential further research.

Acknowledgments

Firstly, I would like to express my deepest respect and gratitude to my supervisor, Prof. Yoichiro Hoshino, Field Science Center for Northern Biosphere, Hokkaido University. His academic rigor has been the greatest support in helping me complete my doctoral research. I remember when I just started my doctoral studies in a new field, he had dedicated half of a year to holding weekly individual seminars with me, despite his busy schedule, to help me quickly master the knowledge related to deep learning. Moreover, he has also provided considerable assistance in not only my studies but also my daily life. From the beginning, when I came to Japan alone, to now, having a child, he has repeatedly extended a helping hand during my most challenging times. I vividly remember all of this. I am also grateful to Prof. Takafumi Gotoh, Assoc. Prof. Koh Nakamura, and Assist. Prof. Toshiyuki Hirata, all from the Field Science Center for Northern Biosphere, Hokkaido University, for their constructive reviews of this study. I am truly grateful for the fellowship provided by JST SPRING, Grant Number JPMJSP2119, which has provided me with financial support. I wish to also thank all the wonderful friends I have made in Japan, as they have made my life rich and vibrant. Finally, I truly appreciate my wife and my parents, for their unparalleled support.

要約 (Summary in Japanese)

ボタン (*Paeonia suffruticosa*) は観賞用の園芸作物としての価値が高く、新しいボタンの品種改良は通常、交配と形質評価をもとにした選抜を行う。多様な遺伝資源を交配に用いた結果、稔性が低下する傾向にあることが知られている。花粉稔性の可視化は育種において重要で、交配の成否を左右する重要な要素となっている。

花粉稔性を可視化する方法について、さまざまなものが開発されておりそれぞれの特徴がある。例えば、*in vivo* 培養法は観察が困難であり、染色による方法は花粉細胞の構造を損なう可能性があり、その後の分析に影響を与えることがある。IFC 法は専門的で高価な設備を必要とする。*in vitro* 培養法と手動によるカウントの組み合わせは非常に時間がかかる。これらの課題に対処するため、本研究では、これらの方法に代わる非破壊的的なディープラーニングベースの二つのシステムを提案した。

第一のシステムは、Mask R-CNN モデルに基づいて構築した。提案したモデルは、花粉粒および花粉管のインスタンス分割に優れており、顕微鏡画像中で個々の花粉粒および花粉管の精密な検出とカウントが可能である。合成データセットをトレーニングデータセットに追加することにより、モデルは花粉管の交差という複雑なシナリオを処理できるようになった。さらに、ポストプロセッシングコードの追加により、モデルは花粉管マスクのスケルトン化を行い、花粉の発芽頻度を計算するだけでなく、花粉管の平均長さを推定することが可能である。この強化により、花粉の発芽性についての洞察が深まった。提案された Mask R-CNN モデルの汎用性は木立牡丹の花粉に限らず、他の植物種にも適用可

能である。さらに、この適応性は転移学習のプロセスを通じて強化でき、ボタンの花粉で訓練されたモデルは、異なる外見を持つ他の植物種の花粉にも適応させることが可能である。この特性は、再構成なしに広範な研究者が花粉稔性の可視性を評価するための貴重なツールとしてのモデルの潜在能力を示している。

第二のシステムは、効率と速度で知られる物体検出モデルである YOLOv7 に基づいている。このモデルは、吸収した花粉の形態的特徴に基づいて花粉の発芽可能性を予測するために使用された。吸水した花粉粒の画像を撮影することにより、サイズ、形状、色などの主要な形態的特徴を分析し、花粉稔性を評価した。安定した花粉の発芽と一致する最適なタイミングで画像を撮影することで、モデルのトレーニングに使用されるデータが正確で関連性のあるものとなる。研究結果は、吸水した花粉粒がその発芽性と相関する特徴を示していた。さらに、提案されたモデルは、これらの特徴に基づいて花粉粒を発芽可能なものとそうでないものの二つのカテゴリに効果的に分類した。YOLOv7 の速度はリアルタイム処理を可能にし、時間に敏感な実験に適している。しかし、今回の研究で使用了データセットに特化している YOLOv7 の特異性を考慮すると、異なる植物種での広範な応用のためには、追加のトレーニングと調整が必要であることが示唆される。この技術的進展は、ボタンまたは他の植物の花粉稔性の迅速な評価において新規的かつ有用なものであると考えられる。