



HOKKAIDO UNIVERSITY

Title	Development of a system for evaluating <i>Paeonia suffruticosa</i> pollen viability using deep learning techniques [an abstract of entire text]
Author(s)	ZHANG, Mengwei
Description	この博士論文全文の閲覧方法については、以下のサイトをご参照ください。 https://www.lib.hokudai.ac.jp/dissertations/copy-guides/
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
Doc URL	https://hdl.handle.net/2115/93480
Type	doctoral thesis
File Information	ZHANG_Mengwei_summary.pdf



学位論文内容の要約/Dissertation Summary

博士（環境科学） 氏名/Applicant ZHANG Mengwei（張 孟偉）

学位論文題名/Title of Dissertation

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）

1. General introduction

Paeonia suffruticosa, commonly known as the tree peony, is a popular ornamental shrub. Until now, various tree peony cultivars have been produced by cross-pollination and artificial selection. Breeding new tree peony cultivars typically involves hybridization and selecting seedlings 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. As seen in other plants, high pollen viability correlates with successful seed production. Various methods are used to assess pollen viability, each with its limitations. For instance, staining methods damage the integrity of the pollen cell, preventing subsequent analysis. The impedance flow cytometry (IFC) method requires specialized and costly equipment. In vitro pollen culture is useful because it can evaluate the potential to retain pollen germination ability. However, the in vitro culture method combined with manual counting of pollen tubes is extremely time-consuming.

Deep learning is a new method of artificial intelligence that 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. To address the above issues, this study proposed two deep learning-based systems as non-invasive alternatives to these methods. 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.

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

The first system was built upon the Mask region convolutional neural network (R-CNN) model (Fig. 1). 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 (Fig. 2). By adding a synthetic dataset to the training dataset, the model can handle more complex scenarios of pollen tube crossing. Additionally, the inclusion of post-processing code allows the model to skeletonize pollen tube masks, enabling it to calculate the pollen germination frequency and estimate 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 mAP₇₀ and 0.633 for mAP@ [0.5:0.95]. In terms of accuracy, the model achieved 0.997 for pollen grains and 0.903 for pollen tubes at the IoU threshold of 50%. These metrics highlight the model's effectiveness in accurately segmenting pollen grains and tubes, confirming its reliability and robustness in pollen analysis. The generalizability of the proposed Mask R-CNN model extends beyond tree peony pollen. Its robust segmentation capabilities make it applicable 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 distinctly different appearances. This feature highlights the model's potential to serve as a valuable tool in diverse research, enabling researchers to assess pollen viability without requiring extensive reconfiguration.

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

The second system was based on the You only look once (YOLO) v7, 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 (Fig. 3). 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, the pollen grains of tree peony with an area greater than 14 μm^2 typically exhibit germinability, whereas pollen grains with less than 12 μm^2 generally do not exhibit such ability. Correlation analysis indicates 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), exhibits a moderate negative correlation with germinability ($\rho=-0.40$). Pollen grains that expand into a circular shape upon hydration typically demonstrate higher germinability. Overall, the findings suggest that the hydrated pollen grains exhibit distinct features correlated with their germinability. Pollen grains with larger areas, brighter colors, and shapes closer to circular are more likely to germinate upon hydration, highlighting the significance of these morphological characteristics as indicators of pollen viability. The proposed model effectively categorized pollen grains into two categories based on these characteristics: germinability and those that do not (Fig. 4). The model achieved a high mAP of 0.979 and a low validation loss of 0.057 with a 50% IoU threshold after training for 800 epochs. Additionally, the precision-recall curves for both classes exhibit high precision and recall, and the AUC-PR is notably high, indicating the model's effectiveness in accurately identifying instances while minimizing false positives. These results all support the strong model performance. However, the model's specificity to the dataset from this study suggests that further training and adjustments would be necessary to apply it broadly across different plant species. The speed of YOLOv7 allows for real-time processing, making it suitable for time-sensitive experiments.

4. General discussion

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. The deep learning-based image detection method is a non-invasive method, which facilitating the subsequent utilization of the detected pollen.

The Mask R-CNN and YOLO models belong to different model types. Mask R-CNN is an instance segmentation model, whereas YOLO is an object detection model, leading to logical differences in the experimental workflow. The Mask R-CNN model 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. The proposed YOLOv7 model can determine the germination potential of hydrated pollen grains based on their

morphology. A YOLOv5 model in a related study was used 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.

The execution time of the Mask R-CNN and YOLO models vary significantly due to their varying architectures. 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. 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. 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.

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

Elucidation for the function of pollen in plant reproduction is expected to 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 further research and potentially enhance breeding strategies.

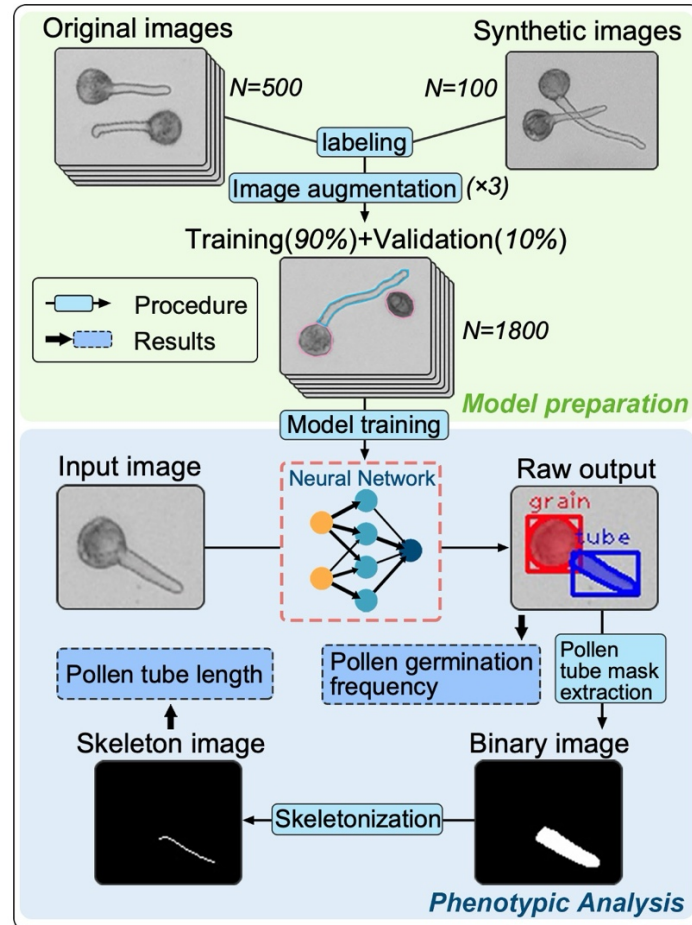


Fig. 1. Workflow and the working principle of the proposed Mask R-CNN model. In total, 500 original and 100 synthetic images were prepared with a size of 692×520 pixels at 300 dpi. A total of 600 images were manually labeled and subjected to data augmentation to generate 1800 images, of which 1620 contained 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. 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. The medial axis algorithm 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.

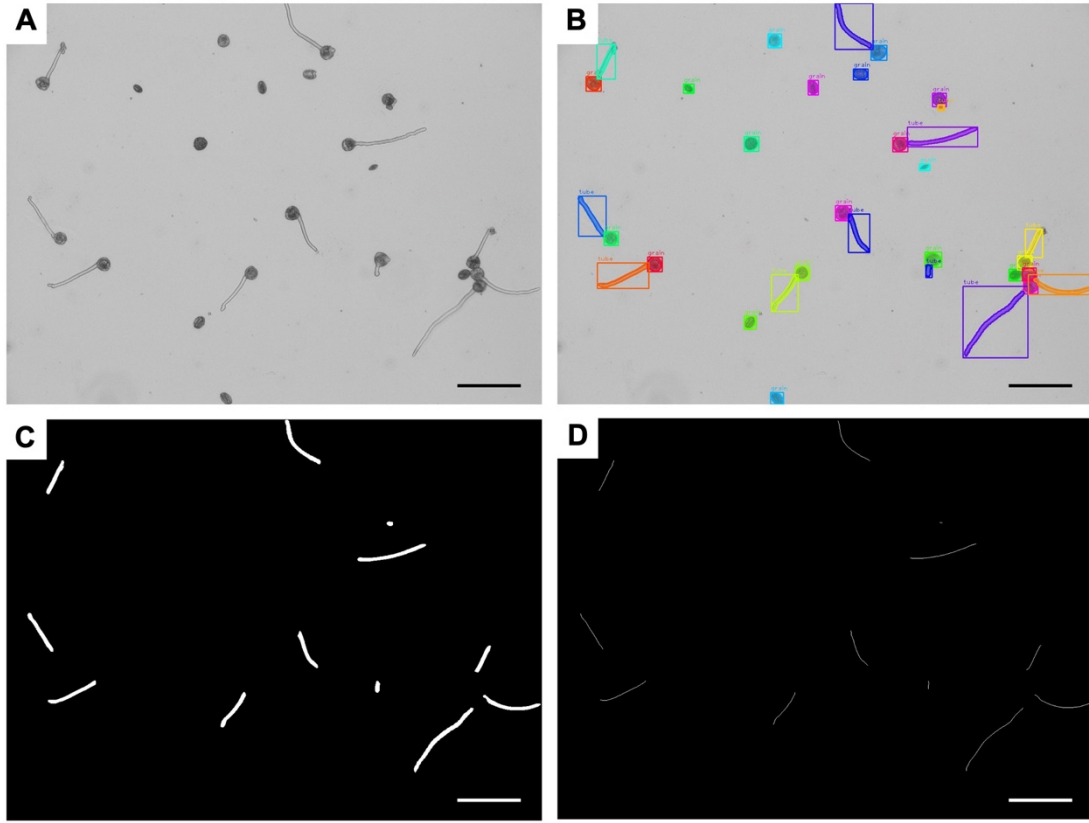


Fig. 2. Output of the proposed model compared with the input image. (A) Input image, (B) raw output, (C) binary mask of pollen tubes, and (D) skeleton of pollen tubes. Scale bars = 200 μm .

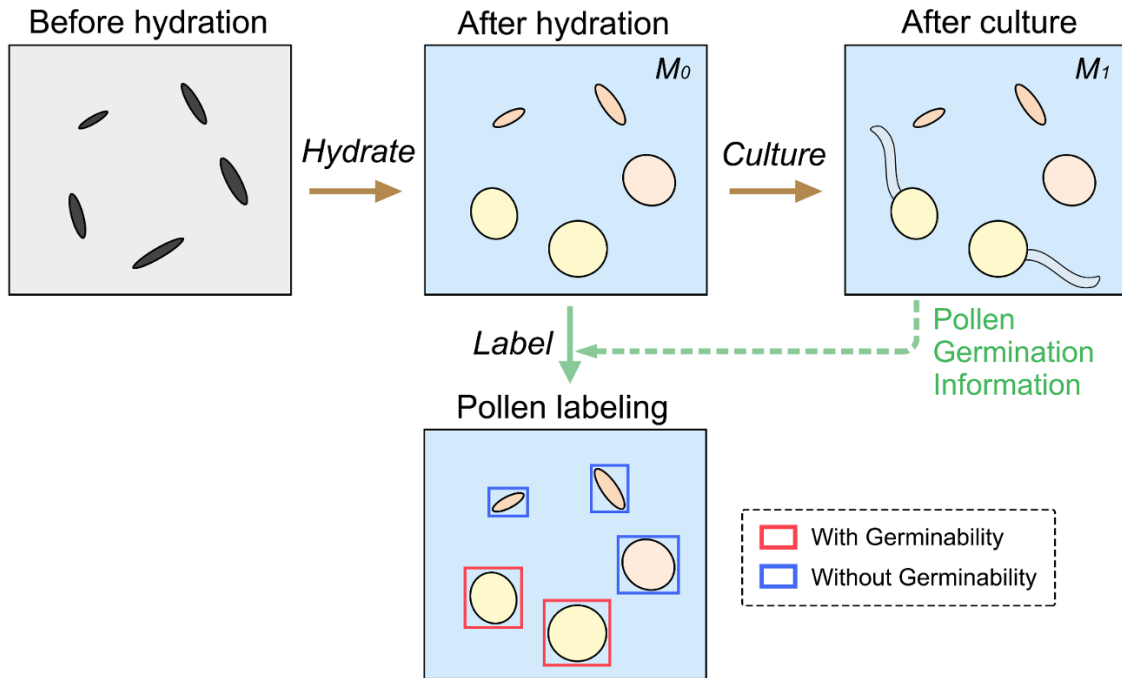


Fig. 3. Image acquisition and labeling for the proposed model. Image acquisition and labeling was performed as follows: during in vitro pollen culture, an initial image, M_0 , was captured once

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 .

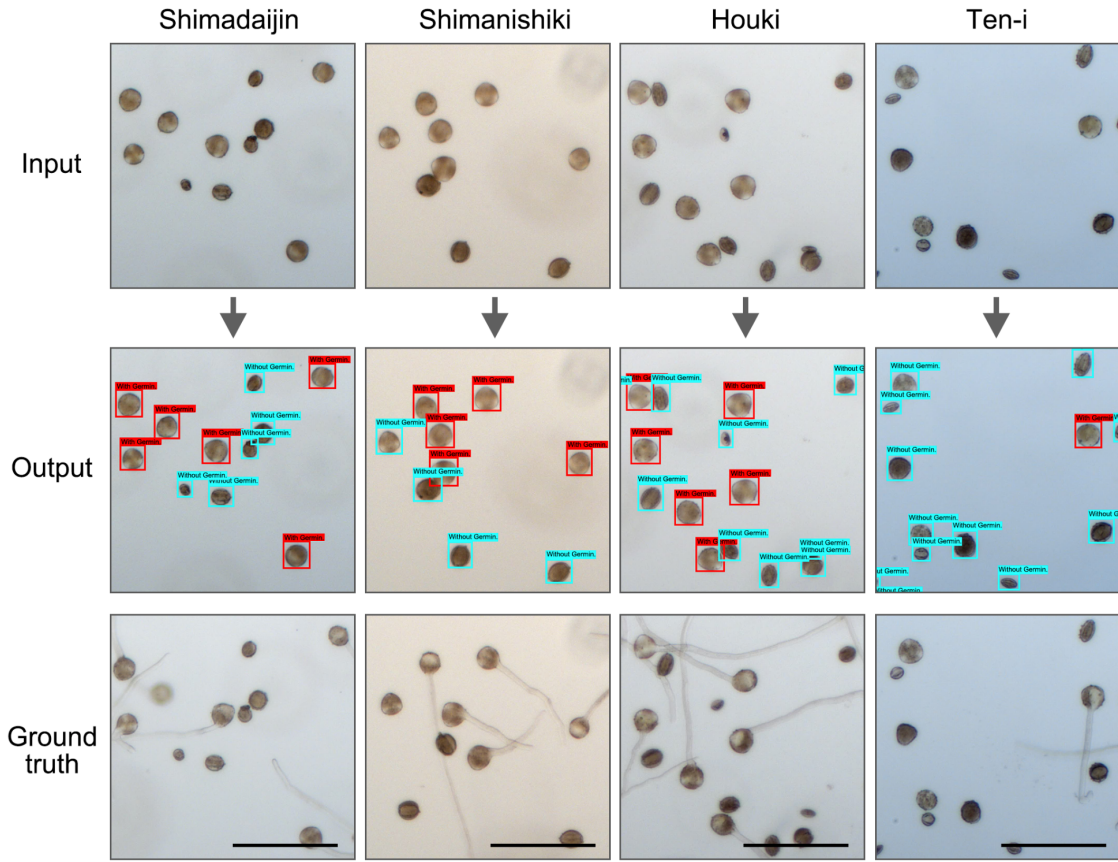


Fig. 4. YOLO model predictions and ground truth for four cultivars of tree peony. The red boxes indicate pollen grains with germinability, whereas the blue boxes indicate those without germinability. Scale bars=1 mm.