Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

From Image to Trait: High-Throughput Phenotyping Pipeline of Early Developmental Stages of Kelp

Sanne Put
Wageningen University & Research
Harald Holm
Wageningen University & Research
Liam Jemison
University of Wisconsin-Milwaukee
Gabriel Montecinos Arismendi
Hortimare BV
Luisa M. Trindade
Wageningen University & Research
Ricardo Silva Torres
Wageningen University & Research
Tijs Ketelaar
tijs.ketelaar@wur.nl

Wageningen University & Research

Method Article

Keywords: seaweed, Saccharina latissima, microscopic image analysis, instance segmentation, deep learning, YOLOv11

Posted Date: November 5th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11348100/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

From Image to Trait: High-Throughput Phenotyping Pipeline of Early Developmental Stages of Kelp

Authors: Sanne Put[†1], Harald Holm[†2], Liam Jemison[3], Gabriel Montecinos Arismendi[4], Luisa M. Trindade[1], Ricardo da Silva Torres[5], Tijs Ketelaar[2*]

[1]Plant Breeding, Wageningen University & Research, Droevendaalsesteeg 1, 6708PB Wageningen, The Netherlands

[2]Laboratory of Cell and Developmental Biology, Wageningen University & Research, Droevendaalsesteeg 1, 6708PB Wageningen, The Netherlands

[3]Department of Mathematical Sciences, University of Wisconsin-Milwaukee, 3200 N Cramer Street, 53201 Milwaukee, Wisconsin, USA

[4]Hortimare BV, Altonstraat 25A, 1704CC Heerhugowaard, The Netherlands

[5]Artificial Intelligence Group, Wageningen University & Research, Droevendaalsesteeg 1, 6708PB Wageningen, Netherlands

[†]These authors contributed equally to this work.

*Corresponding author. Laboratory of Cell and Developmental Biology, Wageningen University & Research, Wageningen, 6708 PB, P.O. Box 386, The Netherlands. E-mail: tijs.ketelaar@wur.nl

Keywords: seaweed, *Saccharina latissima*, microscopic image analysis, instance segmentation, deep learning, YOLOv11

Acknowledgements

We would like to thank the Wageningen Light Microscopy Centre (WU) for the use of their facilities.

Author contributions

S. Put: Conceptualization, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – Original draft. H. Holm: Conceptualization, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – Original draft. L. Jemison: Conceptualization, Investigation, Methodology, Software, , Validation, Writing – Original draft, Writing – Review & Editing, G.M. Arismendi: Conceptualization, Investigation, Writing – Review & Editing , L.M. Trindade: Supervision, Writing – Review & Editing, R. da Silva Torres: Conceptualization, Methodology, Writing – Original draft, Writing – Review & Editing, T. Ketelaar: Investigation, Funding acquisition, Project administration, Resources, Supervision, Writing – Review & Editing.

Funding

This work was supported by the project Seaweed Attachment onto Substrates and Economic Embedding in Dutch Society (SEASEEDS) with file number KICH1.LWV01.20.004 of the research programme Knowledge and Innovation Covenant 2020-2023 Aquatic food production and is financed by the Dutch Research Council (NWO).

Data and code availability

The data and code underlying this article is available at https://github.com/SannePut/From-Image-to-Trait.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process.

During the preparation of this work the authors(s) used ChatGPT in order to improve writing, R code and Python code debugging. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

1. Abstract

Seaweed is gaining attention for its potential as sustainable biomass. However, limited breeding efforts and research on crop management have been performed on seaweed compared to agricultural crops. Successful early development is critical for consistent yields, but high-throughput tools to study this microscopic life stage are lacking. Here we developed a high-throughput phenotyping pipeline for assessment of juvenile stages of kelp including semi-automated image acquisition and sporophyte trait by instance segmentation. We created a single large dataset comprising of 1129 images and 2823 instances which were reclassified to train four different models. Three sporophyte models were trained with different developmental stages to assess the effect of dataset composition on effectiveness and to provide flexibility for users with images of seaweed at different developmental stages. Our models achieve extremely high mean average precision scores (0.872 mAP50 and 0.681 mAP50-95) on mask detection, demonstrating excellent localization of juvenile sporophytes, even when overlapping at high sporophyte density. To follow growth and development or test treatment effects, the *Alignment* model was trained for alignment of images acquired sequentially over time, which achieved an mAP50 of 0.994 and mAP50-95 of 0.943 on masks. We compared the effectiveness and efficiency of the model with three experts. The model effectiveness was comparable to the experts, but 18x faster. This study provides an accurate, high-throughput phenotyping method for faster development of seaweed research that excludes inter-and intra-personal variability during data assessment. Potential applications are testing the effect of (a)biotic stresses, breeding programs and quality control of seeding material.

2. Introduction

Seaweed is increasingly recognised as a sustainable biomass source by multiple stakeholders (Cai et al., 2021; European Commission, 2020; Vazquex Calderon et al., 2022). As production of terrestrial agricultural biomass puts increasing pressure on natural resources and soil quality, seaweed production has the potential to partially substitute agricultural biomass for several applications. Compared to terrestrial agriculture, seaweed aquaculture generally requires less resources and can provide ecosystem services, including nutrient removal and an increase in local aquatic biodiversity (Chopin et al., 2001; Duarte et al., 2017; Eswar et al., 2025). Seaweed is a multi-purpose biomass with applications in a wide range of industries, including food, feed, cosmetics and bioplastics (Holdt & Kraan, 2011; Polat et al., 2023). Although seaweed has a wide range of applications, European seaweed farming remains at its infancy, particularly due to high labour and investment costs compared to wild harvesting (Zaat, unpublished). To avoid decline of natural seaweed populations due to overharvesting, seaweed farming should complement wild harvest (Araújo et al., 2021; Buschmann et al., 2017; Rebours et al., 2014). Compared to terrestrial agricultural crops, seaweed remains understudied, even though new tools and protocols are essential for upscaling of seaweed farming (Charrier et al., 2017).

In this study, we focus on seeding material of the kelp *Saccharina latissima*. Among cultivated kelps in Europe, *S. latissima* is an important species for seaweed farming due to its fast growth and adaptation to environmental conditions (Bak et al., 2018; Kraan, 2013; et al., ; Van Der Molen et al., 2018). To obtain high and stable yields, proper establishment of seaweed seeding material is crucial (Charrier et al., 2017). Seeding material consists of juvenile stages of kelp and its production requires working with its diplohaplontic life cycle (Theodorou et al., 2021; Visch et al., 2019). Seeding material consists of a mix of microscopic haploid, dioecious gametophytes and juvenile diploid sporophytes that eventually develop into the macroscopic kelps. Gametophytes can be cultivated vegetatively or under conditions that induce the development of egg cells on female and spermatozoids on male gametophytes, respectively. Following fertilisation, the sporophyte originates from a zygote that initially develops associated with the female gametophyte and subsequently develops highly elongated, adhesive secreting cells referred to as rhizoids

from basal side to independently attach to a substrate (Theodorou et al., 2026). Multiple egg cells can develop from a single gametophyte. After fertilization, this results in a “bouquet” of sporophytes, all connected to a single female gametophyte. Two processes during early sporophyte development are essential for successful establishment on a farm (Charrier et al., 2017). First, high fertilisation success increases the number of sporophytes produced per gametophytic biomass. This is highly relevant since vegetative bulking of gametophytes is a slow process (Huang et al., 2022; Veenhof et al., 2025). Higher fertilization rates increase the number of juvenile sporophytes available for seeding. Second, successful attachment to a substrate is essential for subsequent sporophyte development and growth (Theodorou et al., 2026). A complicating factor in studying this life stage transition and early growth is its microscopic scale, as the diameter of egg cells and zygotes is smaller than 40 µm (Theodorou et al., 2021) and lengths of the developmental stages used in seeding material range from 50 to 500 µm (Umanzor et al., 2020).

To study the effect of genetic and environmental factors on early sporophyte development, large numbers of genotypes and environmental conditions must be evaluated. Phenotyping of the relevant developmental stages relies on labour intensive microscopy. A high-throughput phenotyping pipeline from image acquisition to trait extraction to efficiently study early development of *S. latissima* could overcome this bottleneck. Potential traits to speed up image acquisition and reduce manual labour are automated scanning and stitching of multi-well plates. Once images have been collected, multiple tools are available for data analysis that have potential for automation. A commonly used method in automated microscopy image analysis is thresholding (Bangare et al., 2015; Fang & Zhong, 2023; Nichele et al., 2020; X. Yang et al., 2001). Thresholding is a simple and computationally efficient method that separates foreground from background based on pixel intensities and is therefore suitable for images with large contrast (Mohan & Chandar, 2025). However, thresholding has two important limitations. First, thresholding cannot differentiate between different types of objects if pixel intensities show too much overlap. In many applications like ours this is difficult to guarantee. Second, connected or overlapping objects will be identified as a single object instead of two separate objects. These limitations are particularly problematic for early seaweed developmental stages because sporophytes can overlap and remain connected by the female gametophytes, requiring more advanced image segmentation approaches.

Deep learning-based segmentation has become an important tool for automated trait extraction in agricultural research because it can detect and characterise objects in complex images (Lei et al., 2024; Luo et al., 2024). Deep learning is a type of machine learning that uses multi-layer neural networks trained on annotated image data. Segmentation methods for image analysis can be classified as semantic, instance and panoptic segmentation (Mohan & Chandar, 2025). Semantic segmentation assigns a class to each pixel but does not distinguish individual objects, whereas instance segmentation identifies individual objects of the same class. Panoptic segmentation combines both approaches but cannot distinguish overlapping instances. We treat sporophyte identification as an instance segmentation problem since sporophytes tend to overlap within bouquets and at high densities, requiring prediction of separate masks for each detected object (Mohan & Chandar, 2025). We selected a state-of-the-art one-stage detector, YOLOv11-seg (Jocher & Qiu, 2024), because of segmentation effectiveness (performance), computational efficiency and accessibility for researchers with limited deep-learning experience.

Here, we developed a high-throughput phenotyping pipeline for efficient screening of *S. latissima* seeding material in large scale experiments. This pipeline integrates automated image acquisition, image alignment and deep-learning-based instance segmentation to extract phenotypic traits from microscopic images. We evaluated the effectiveness and efficiency of multiple sporophyte instance segmentation models across different dataset compositions and an automated image alignment method to follow longitudinal growth and development and comparison between treatments. Using the state-of-the-art model YOLOv11-seg we developed a user-friendly pipeline that requires limited computational resources, while balancing efficiency and effectiveness. The datasets are available at

https://github.com/SannePut/From-Image-to-Trait. By reducing the time and labour required for microscopic phenotyping of kelp seeding material, this pipeline allows for screening of large genetic diversity and testing responses to (a)biotic stresses, with the potential to enable faster method development and breeding progress.

3. Materials and methods

3.1. Kelp material generation

Gametophyte cultures were isolated from individual kelps and provided by Hortimare B.V. (Heerhugowaard, Netherlands). The gametophyte cultures were maintained in 34 ppt salinity artificial seawater (Instant Ocean, Sarrebourg, France) with F/2 nutrients without Fe (Guillard & Lorenzen, 1972) at 11°C and 8 µmol photons m−2 s −1 red light and a 14-hour photoperiod (Theodorou et al., 2021). Before making crosses, female and male gametophytes were separately fragmented with Eppendorf micro-pestles (Eppendorf, Hamburg, Germany) and combined at a 1:1 ratio. Gametogenesis was induced by cultivating the combined cultures in 34 ppt salinity artificial seawater (Instance Ocean, Sarrebourg, France) with F/2 nutrients (Guillard & Lorenzen, 1972) at 11°C and 25 µmol photons $m^{-2}s^{-1}$ white light and a 14-hour photoperiod (Theodorou et al., 2021). Either the combining of the gametophytes occurred in the six well plate or on the coverslip, or the culture was detached with an inoculation loop and then seeded onto a coverslip or in a six well plate.

3.2. Data acquisition

An Andor-revolution Spinning Disk microscope (Oxford Instruments, Abingdon, UK), using a Nikon Plan Fluor 4x/0.13 numerical aperture (NA) objective (Amstelveen, the Netherlands) and black-and-white Andor iXON888 EM-CCD camera (Oxford Instruments, Abingdon, UK), controlled by Metamorph (version 7.10.2.240, BioVision Technologies, Exton, PA, USA) was used to acquire 993x993 pixel images of kelp seeding material seeded on coverslips placed in 3D printed holders (Supplementary design 1) in six well plates or on the bottom of the six well plates. We imaged the coverslip through a coordinate generating Python script (Supplementary script 1) that allows to image a large region of the coverslip in a standardised manner. This python script accounts for gradual differences in the z-axis. The resulting individual images are then stitched in Fiji (ImageJ 1.53c; National Institutes of Health, USA) through an ImageJ macro (Supplementary script 2).

3.3. Model training

We used for model training the one-stage detector YOLO for instance segmentation predictions for its computational efficiency, while maintaining detection accuracy comparable to two stage detectors (Jocher & Qiu, 2024). Additionally, its extensive documentation allows for ease of use. Experiments were conducted within a Python environment using YOLOv11 from the Ultralytics package (2024). The computer configuration used for the experimentation is: Intel i5-8500 CPU, 32 GB of RAM, GeForce RTX 4060 Ti 16 GB GPU and Windows 10 Enterprise OS 17763.8511. The segmentation YOLO models were pre-trained on the COCO dataset (Lin et al., 2014). Our large dataset is made up of images with associated labels put in certain classes. Labels are regions within the images that show an object. A class is a label that adheres to determined morphological restrictions to categorize objects within images and indicated with “”. A labelled region with a class associated in an image is defined as the ground truth. The dataset was reclassified based on the training goal after which the models were fine-tuned using the resulting different versions of the datasets. These different versions of the models are indicated by *italics*. We compared the effectiveness of the model sizes n, s, m, and l, which correspond to the number of parameters used by the model to predict an object. Predictions come in two forms in YOLO, boxes and masks. Boxes are the minimum square that can contain the object. Masks are the association of the pixels with an individual object of a class. We also enabled class agnostic non-maximal suppression (*agnostic_nms*), which merges overlapping recognitions. The datasets were split up into a training, validation, and test sets containing, respectively, 70%, 10%, and 20% of the datasets. We trained for 3000 epochs using an image size of 993x993 pixels. *agnostic_nms* was used on the resulting bounding box to merge highly overlapping detections. A patience of 400 was used for early stopping of the training. Due to memory constrictions the batch size varied between the different sizes of models. For YOLOv11n-seg, s-seg, m-seg, and l-seg, we

used, respectively, a batch size of 16, 14, 10 and 8. Other training settings, such as the learning rate schedule, data augmentation and optimizer were kept at the default setting of YOLO as provided in the Ultralytics package (Jocher & Qiu, 2024).

3.4. Model testing

For validation of the model performance, we use multiple parameters to look at its efficiency and effectiveness. We define efficiency in time, with high efficiency taking less time than compared to human labour. This represented by the timing of actions as well as the inference time per image. Effectiveness is the accuracy of the prediction or measurement. Proxies for effectiveness are Precision, Recall, F1, mean Average Precision 50 (mAP50), mean Average Precision 50-95 (mAP50-95) and comparison with the Ground Truth. Precision is the proportion of true positives of all positive predictions, representing the likelihood that a sample in the test set is correctly identified by the model. Recall quantifies the proportion of true positives among all actual positives, assessing the capability to all instances of a class. F1 represents the harmonic mean of precision and recall, which gives an assessment of the model performance taking both false positives and false negative into account. Average Precision (AP) determines the area under the precision-recall curve, summarizing the model's precision and recall performance into a single value. mAP50 is the average AP values across multiple classes at Intersection over Union (IoU) (Fig.S1) of 0.5. mAP50-95 is the average of the average AP values of multiple classes at IoU thresholds from 0.5 to 0.95 with steps of 0.05. To test the efficiency and effectiveness of the instance segmentation vs human annotation, we asked three experts to annotate sporophytes within a representative large image and time the amount of time it takes to annotate the image. The model that had the highest mAP50-95, was used for inference on the same image. The resulting regions of interest (ROI) were counted and measured and compared to the Ground Truth annotations to determine the effectiveness of the experts and the model. Timing of the actions was used to determine the efficiency of the model compared to experts.

3.5. Detection and post-processing

Processing of large images leads to constraints in the vRAM capacity of GPUs. To preserve the high resolution of the images we setup a pipeline for detection and post-processing of splitting large images into smaller images and reassembling the instance into the larger image. For obtaining optimal results, there are number of parameters that can be tweaked for the use case. For the prediction the Ultralytics package includes the confidence & IoU thresholds (Fig.S1). Instance post-processing can be adjusted by the variables n_splits, border_threshold, overlap_threshold, threshold, min_mask_size (Table 1).

In our post-processing pipeline, we developed a new merging strategy. To measure its effectiveness against the other merging strategies non-maximal suppression (NMS) and non-maximal merging (NMM), as well as no merging, we compared merging strategy performance on the same representative large image used for the comparison with experts.

Table 1 Parameters and their default settings for detection and post-processing of instances.

Parameters	**Default**	**Definition**
Confidence	0.2	Certainty of the prediction by the model
IoU	0.7	Ratio of the overlap area of the prediction with the ground truth and the merge area of the prediction with the ground
N_splits	4	Due to vRAM constraints large images might not be able to be processed. This variable allows for the larger image to be split up into smaller images after which the larger image is reassembled with detections in the correct position. Detection of large individuals (>2mm) is improved with lower number of splits.
Border_threshold	25	Sporophytes can end up on the border of a split or on the edges of the acquisition region. To prevent incomplete detections, we

		determine how close the detections are to which border of the split and place it in a list to be merged with the corresponding detection. This closeness to the border can be tweaked with border_threshold.
Overlap_threshold	0.6	If two detections overlap more than overlap_threshold, then these detections are merged.
Threshold	100	The detections identified due to their proximity to the border of the split, are checked whether they can merge with a corresponding detection in the neighbour split. Threshold determines the distance how close the detections need to be to each other to merge.
Min_mask_size	400	Min_mask_size allows for the filtering out of detections below a certain size.

4. Results

4.1. Automated and efficient phenotyping pipeline

To reduce time and manual labour required for microscopic phenotyping of kelp juveniles, we developed a high-throughput phenotyping pipeline that combines microscopic image acquisition and automated trait extraction (Fig.1). This pipeline allows for large scale experiments that are time and cost effective. The first step includes image acquisition of a user-defined region within a well of a multiple-wells plate. Depending on the magnification and size of this region, the acquisition can consist of multiple images (tiles) captured by the camera and subsequently stitched together. A two-dimensional tilt correction was applied to correct for small deviations in sample placement from the horizontal axis. Tilt corrections were performed by adjusting the z-position in the corners of the coverslip and assuming a linear two-dimensional z-height gradient over the coverslip. The image analysis pipeline consists of two components. The first component, the *Alignment* model, enables automatic alignment of images acquired sequentially over time to track growth and development or test treatment effects. The model detects circular reference markers drawn on the coverslip with a permanent marker in two corners. Recognition of the same markers on several images allows for alignment of images, based on the coordinates of the markers. The second component is a model for automated detection and trait extraction of individual sporophytes of which we developed three different version. As processing of large images leads to constraints in vRAM capacity of GPUs, the large image is split into smaller images with overlapping regions for detection. A post-processing pipeline reassembles the detections from sporophytes that cover multiple splits into one detection on the larger image. The duration of the whole pipeline depends on different factors, such as the magnification, camera settings and the number of wells and region to be imaged. To give an indication of the efficiency of the pipeline, we measured the duration of each step in the pipeline using standard settings (4x magnification, ~2.5 cm^2 imaging surface per well for a total of 6 wells). Setting up and scanning the 6 wells takes approximately 23 minutes. Both the stitching and trait extraction are very rapid and can be performed in batches, allowing for automatic analysis overnight. The most time-consuming and laborious part of the process for researchers is the preparation of samples.

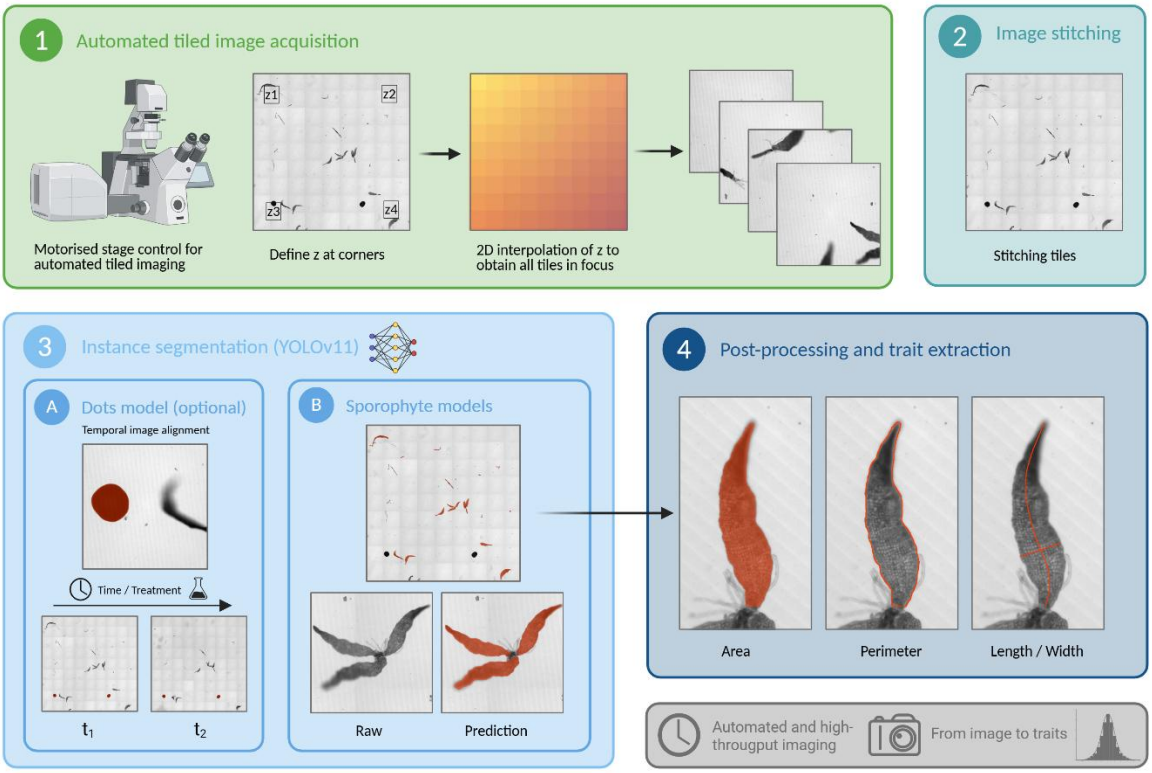


Fig.1 *Schematic overview of the semi-automated high-throughput phenotyping of seaweed material. (1) The pipeline starts with the acquisition of the images. Adjustment of the z-position at the four corners of the imaging area allows for are two-dimensional interpolation of z-positions over the whole imaging area when the surface is not levelled. (2) Separate images are stitched to the larger image. (3) Two models based on YOLOv11 allows for phenotyping the seaweed. The Alignment model is optional and can be used to align different images taken over time to follow sporophyte growth and development or before and after a treatment. The sporophyte models recognise individual sporophytes. (4) Post-processing reassembles detections from different splits into the larger image and automatically extracts sporophyte traits*

4.2. Large and diverse dataset of early kelp developmental stages

We created a single large dataset with labelled images of early developmental stages of *S. latissima* to extract trait values and markers for image alignment (Fig.2). The dataset consists of 1129 images of 993x993 pixels, which are divided over a training, validation, and test sets with proportions of 70%, 10%, and 20% respectively. The training set was used during training of the model, the validation set was used during training to evaluate the training progress, and the test set allows for independent testing of model performance after training. Images were captured with 4x magnification with an NA of 0.13, allowing for sufficiently high resolution to detect single cells and differences in shape and pigmentation of cells. Stitching multiple images allows for analysis of later developmental stages, whose sizes can exceed the boundaries of a single image, while maintaining high resolution. The dataset contains in total, 2823 class labels, which are enclosed in quotation marks throughout the text (e.g. “>64 cells”). From early to late developmental stage, we included five different sporophyte classes (Fig.2). The egg stage (“Egg”) with single rounded cells on the female gametophytes, the zygote stage (“Zygote”) with single, elongated cells on the female gametophyte and young sporophytes of less than 64 cells (“<64 cells”) and more than 64 cells (“>64 cells”). The class “>64 cells” also includes sporophytes with multiple cell layers. A last sporophyte class, Out-of-focus sporophytes (“OFS”), contains sporophytes that are out of focus because they lie partially or completely outside the focal plane, for example when the blade floats while rhizoids remain attached to the coverslip. These sporophytes are always larger than 64 cells in size. Floating can be remedied by placing a coverslip on the sample to flatten the sample. However, this method is not compatible with non-destructive phenotyping is needed to assess, for example, treatment effects or

growth. A class “Marker” consists of circular markers, drawn with a permanent marker, which can be used for image alignment (Fig.2).

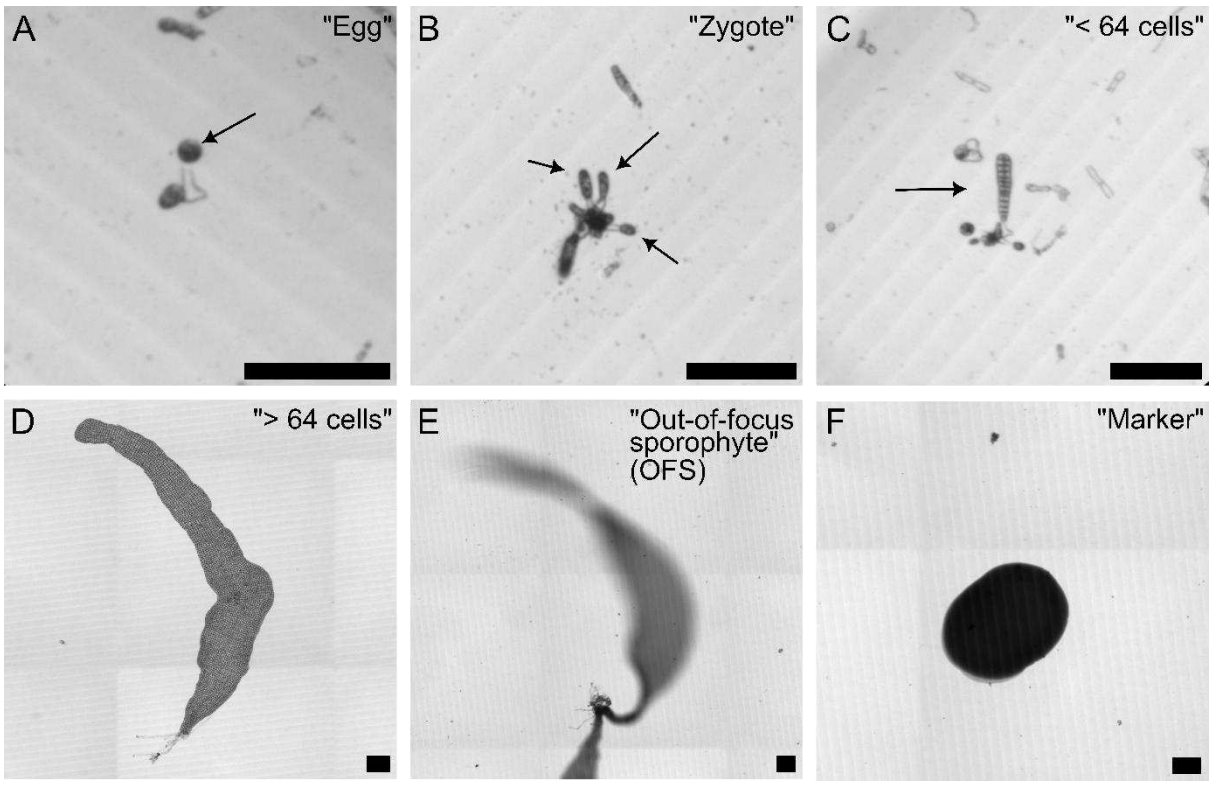


Fig.2 Representative images of the classes within the annotated kelp datasets. A shows a spherical egg cell on the oogonium next to an undifferentiated female gametophyte cell (black arrow). B shows three single celled elongated zygotes indicated by the black arrows. C indicated by the black arrow shows a young sporophyte containing less than 64 cells. D shows an older sporophyte with a large-size and several rhizoids at the basal side. E is a large-size older sporophyte floating upright in the medium. F shows a dot on the coverslip for alignment of global images. Lines on the images are artefacts caused by image acquisition with the spinning disk in the light path. All images are extracted from similar large-scale images. All scale bars represent 200 µm

4.3. Class reclassification enables application-specific model development

To accommodate different applications, the large, annotated dataset was reclassified into four different datasets. These reclassified datasets were used to train different models. Three models included labels for different developmental stages of sporophytes, whereas the fourth model was designed for image alignment. The sporophyte models cover the sporophyte development from induction to the stage that macroscopic analysis can be performed. The alignment was trained to detect circular markers circular markers in images acquired sequentially over time. This enables image alignment and therefore monitoring of longitudinal growth and development and comparison between treatments. The different models are indicated in *italics* throughout the text, and an overview of the datasets is provided in Table 2. The four different models differ in which classes were reclassified and focus on later developmental stages (excluding “Egg” and “Zygote”). The *divided developmental stages* (*DDS*) model was trained with two different classes for “<64 cells” and “>64 cells”, where “OFS” was combined with “>64 cells”. The *all developmental stages* (*ADS*) model combined “<64 cells”, “>64 cells” and “OFS” into a single class. The *>64 cells* model was trained to classify a single class combining “>64 cells” and “OFS”. Finally, the *Alignment* model, which consists of the “Marker” class. In this study we compare the effectiveness, efficiency and limitations of these four different models. The datasets are available on https://github.com/SannePut/From-Image-to-Trait, which allows users to adjust and retrain the current models.

Table 2 Composition of datasets for the four different models. For each model, the number of instances for each class and the number of images is indicated for the training and validation set. (a) The Divided developmental stages (DDS) model consists of "<64 cells" and a combined class for ">64 cells" and "OFS". (b) The All developmental stages (ADS) model consists of one class in which of "<64 cells", ">64 cells" and "OFS" were combined into one class. (c) The >64 cells model consists of one class in which ">64 cells" and "OFS" were combined into one class. (d) The Alignment model consists of the class "Marker".

a

DDS			
Class	**Training**	**Validation**	**Test**
"<64 cells"	295	35	131
">64 cells" + "OFS"	812	121	228
Number of images	609	87	174

b

ADS			
Class	**Training**	**Validation**	**Test**
"<64 cells" + ">64 cells" + "OFS"	1107	156	359
Number of images	609	87	174

c

>64 cells			
Class	**Training**	**Validation**	**Test**
">64 cells" + "OFS"	825	131	207
Number of images	609	87	174

d

Alignment			
Class	**Training**	**Validation**	**Test**
"Marker"	217	31	54
Number of images	609	87	174

4.4. High effectiveness of state-of-the-art models

Effectiveness of the instance segmentation models is crucial when replacing manual work. The four different models we trained all showed high effectiveness, but there are differences between models in terms of performance metrics (Fig.3; Table 3; Fig.S2; Table S1). For the sporophyte instance segmentation models, the effectiveness was higher for the *DDS* than the *ADS* model. The *DDS* model obtained high effectiveness with a mAP50-95 of 0.711 and 0.681 for boxes and masks, respectively. The best performing models for *ADS* were the smaller model architectures with YOLOv11n-seg obtaining a mAP50 of 0.891 and 0.701 for boxes and masks respectively and a mAP50-95 of 0.872 and 0.649 for boxes and masks respectively. On the other hand, for *>64 cells* larger model architectures increase effectiveness. YOLOv11m-seg obtained the highest mAP50 and mAP50-95 for both the boxes (mAP50=0.837 and mAP50-95=0.667) and masks (mAP50=0.853 and mAP50-95=0.643). The *Alignment* model for alignment of images had very high effectiveness with a mAP50-95 score for masks of 0.943. Inference time per image ranges between 0.07 and 0.13 second and mainly depended on the model size, rather than composition of the datasets. Together this shows that our models are efficient and accurate in the prediction of markers and sporophytes.

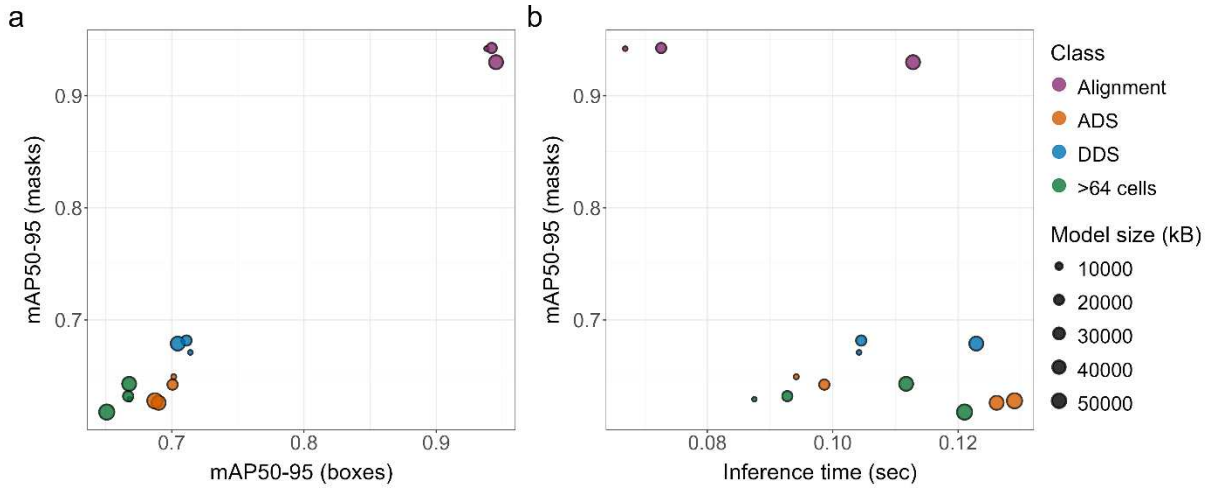


Fig.3 *Mean Average Precision (mAP) for boxes and masks of test set and inference time for different model architecture sizes and four different datasets (>64 cells, All developmental stages (ADS), Divided developmental*

stages (DDS), Alignment). (A) MAP50-95 for boxes and masks. (B) MAP50-95 for masks versus inference time (ms) per image

Table 3 Effectiveness scores for the test set of the best model for each dataset with (a) Divided developmental stages (DDS), (b) All developmental stages (ADS), (c) >64 cells and (d) Alignment. Scores include Precision (True Positives / (True Positives + False Positives)), Recall (Ture Positives / (True Positives + False Negatives)), mean Average Precision at IoU 50 (mAP50), mean Average Precision a 0.05 step increase from IoU 50 to 95 (mAP50-95) and F1(2 x (Precision x Recall) / (Precision + Recall)).

a

DDS	**Box**	**Mask**
Size (kB)	20319	
Precision	0.741	0.733
Recall	0.885	0.877
mAP50	0.855	0.856
mAP50-95	0.711	0.681
F1	0.806	0.798

b

ADS	**Box**	**Mask**
Size (kB)	6251	
Precision	0.908	0.907
Recall	0.827	0.811
mAP50	0.891	0.872
mAP50-95	0.701	0.649
F1	0.866	0.856

c

>64 cells	**Box**	**Mask**
Size (kB)	44359	
Precision	0.808	0.788
Recall	0.872	0.853
mAP50	0.837	0.853
mAP50-95	0.667	0.643
F1	0.839	0.819

d

Alignment	**Box**	**Mask**
Size (kB)	20291	
Precision	0.996	0.996
Recall	0.984	0.981
mAP50	0.994	0.994
mAP50-95	0.942	0.943
F1	0.989	0.989

4.5. Post-processing accurately merges sporophytes over tiles

Deep convolutional neural network models, such as YOLOv11, typically rescale image inputs to a fixed size such as 640x640 pixels. Processing large images such as our high-resolution microscopy stills requires either down sampling; increasing the input size, which substantially increases the computational cost of training and inference; or slicing inputs and performing inference on each slice independently. To maintain detection quality without excessively sacrificing inference time, we used slicing during training and inference.

Slicing typically improves detection of small objects but can lead to extra detections when a large object such as a mature sporophyte is present in multiple slices. Typical approaches to resolving extra detections use bounding box overlap and either NMS (Felzenszwalb et al., 2010; Girshick et al., 2014; Redmon et al., 2016) which effectively resolves overlapping detections in the interior of slices but tends to truncate

detections on the boundaries, or NMM which can effectively resolve detections across slice boundaries but performs poorly in crowded scenes.

Algorithm 1: Algorithm: Merge Within Slice

Data: Detections $M_1, \ldots, M_N$ such that M_i is a binary mask
Result: Merged Detections $D = M_1, \ldots, M_k$
$D \leftarrow \emptyset$;
for $i \leftarrow 1$ **to** N **do**
  $M_i \leftarrow \text{BinaryClosing}(M_i)$;
  $M_i \leftarrow \text{BinaryFillHoles}(M_i)$
  for $M \in D$ **do**
    $I = \text{CalculateMaskIOUInBBoxIntersection}(M, M_i)$;
    if $I > t$ **then**
      $M \leftarrow \text{Merge}(M, M_i)$;
    end
  end
end

Fig.4 *Algorithm to merge masks within a slice, which can also be used to merge masks within merged slices. When the bounding boxes of two mask detections overlap, the intersecting area is isolated. The intersection over union (IoU) is then calculated of the two masks within the isolated intersecting area. If the IoU is above a determined threshold(t) then the two masks are merged*

We present a novel algorithm for merging detections produced by sliced inference that uses instance mask overlap with one rule for merging detections within a slice, and a second rule for merging detections across slices, henceforth called IoU based Slice-Overlap Merging (ISOM). We implement the algorithm using the Slicing Aided Hyper Inference (SAHI) package (Akyon et al., 2022), and show that this approach achieves significantly better mAP50-95 scores on a validation test than NMM, NMS, or no merging at all. Fig.4 shows our merging algorithm, and Fig.5: Comparison of merging strategies compares the performance of our strategy with NMM, NMS, and no merging, demonstrating clear improvement in mAP50-95 and F1 score.

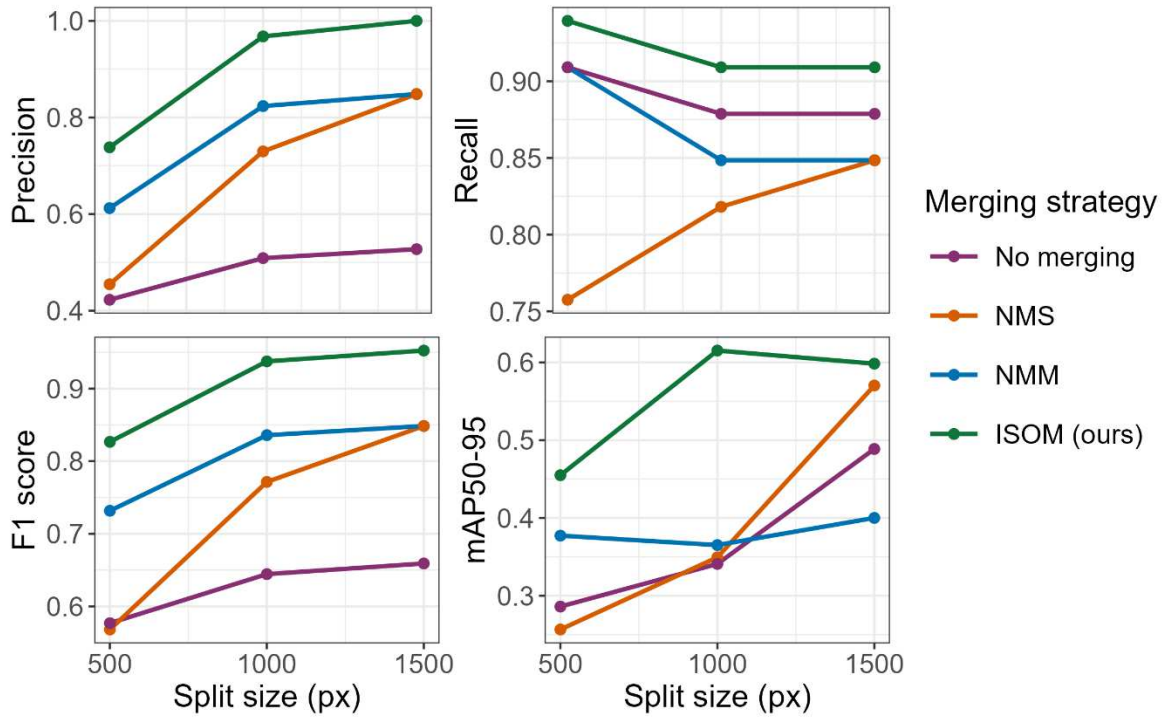


Fig.5 *Comparison of merging strategies NMM (non-maximal merging), NMS (non-maximal suppression), no merging and our strategy in the four metrics: Precision, Recall, F1 and mAP50-95 against increasing split size of the images*

4.6. The model processes images faster and has comparable accuracy to experts

Both accuracy and processing time are important when choosing between manual or automated image analysis. To analyse these aspects, we compared the effectiveness of the model and three experts against the ground truth. The ground truth was annotated by the authors Put and Holm, who also annotated the original dataset. The image contained in total 33 sporophytes, of which 30 were recognised by the model and 26, 27, and 28 respectively by the experts. Both the model as well as the experts did not recognise smaller sporophytes (Fig.S6). The correlation between the ground truth and the model or experts was high, ranging from 0.97 to 1.00 across all traits (area, length, width and perimeter; Fig.6; Fig.S3; Fig.S4; Fig.S5). In terms of efficiency, the model was faster with a processing time of 1 minute compared to a processing time of 10, 20, and 23 minutes for the experts. This means that more data can be analysed using less resources when using automated image analysis compared to manual image analysis.

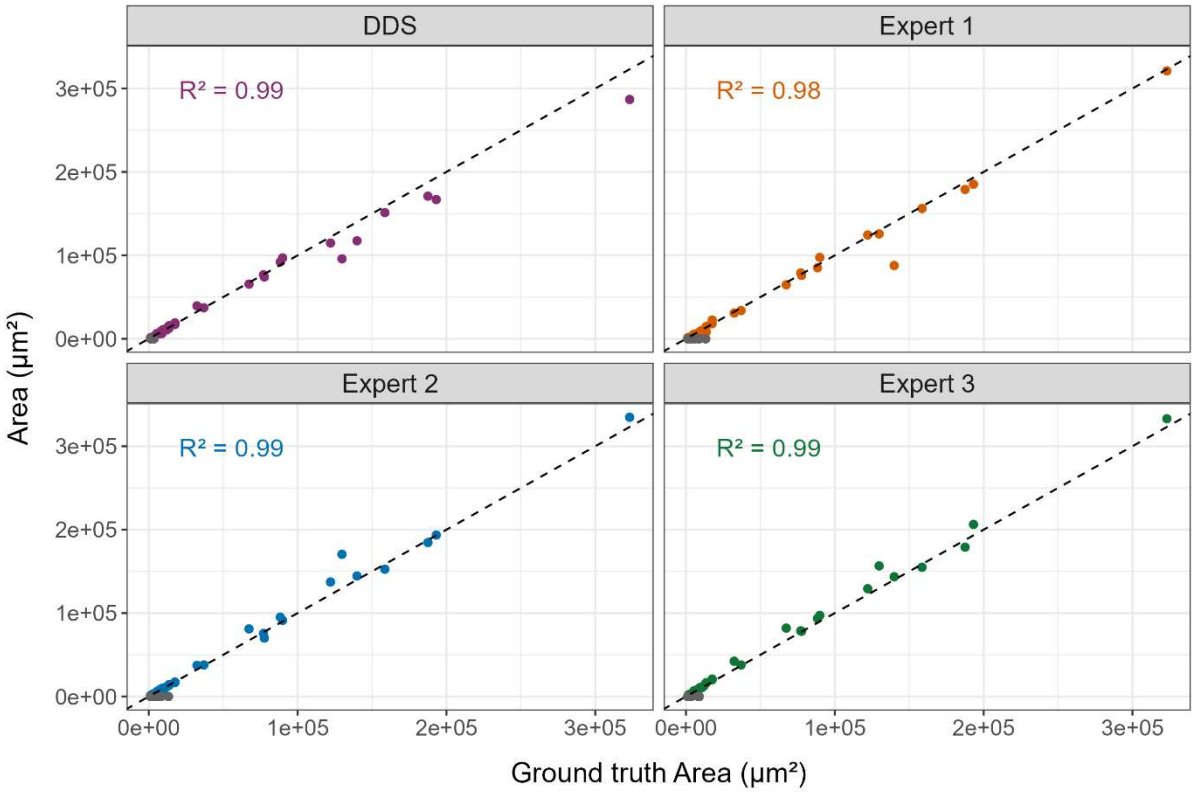


Fig.6 *Comparison of the model (DDS) and three experts with the ground truth for the trait area. Pearson correlations are indicated for the model and each expert. Grey points indicate sporophytes that were not detected by the model or expert and are not included in the calculation of the* R^2

4.7. Limitations

Although the model effectiveness was high, understanding the limitations of instance segmentations models is important for improving models, retraining or adaptation of image acquisition. To determine the nature of our model's limitations, we identified when the model produced inaccurate segmentation results. The cases of failure mainly occurred in images with a high density of sporophytes and larger sporophytes that started to overlap or were partially out of focus. This created difficulty in separating individuals, especially when sporophytes overlap or when the gametophyte was out-of-focus (Fig.7).

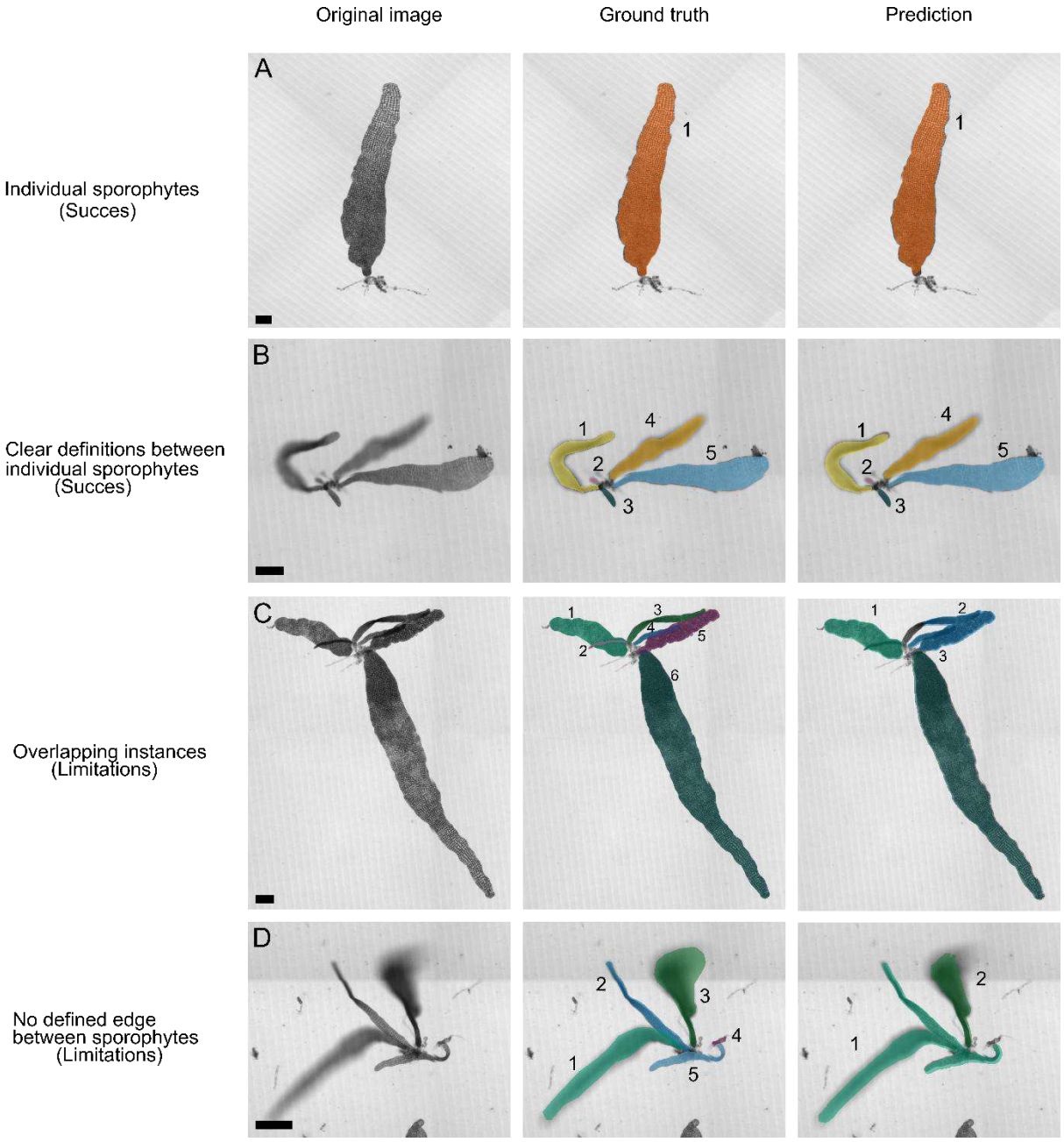


Fig.7 *Limitations and success cases of the detections by the model and merging by the pipeline. The images from left to right are ordered as such: Original image, Ground truth of the classes, and Prediction by the model. A shows an individual sporophyte clearly visible and thus an accurate detection as result. B shows detection when the individual sporophytes of the bouquet are more separate. C shows the instance of overlapping sporophytes causing the detection of a singular sporophyte instead two (for detection 1 & 2) or three (for detection 3,4 & 5) different individuals. (D) When there is no defining edge between the sporophytes in a bouquet, such as for detection 1, 2 and 5, the model also predicts a single sporophyte instead of three different individuals*

5. Discussion

5.1. Affordable high-throughput phenotyping pipeline as basis for future development

In this study we provide a pipeline for high-throughput data acquisition and analysis of microscopic kelp seeding material usable for research and commercial applications of seaweed. This pipeline allows for large-scale acquisition and analysis of high-resolution microscopic images reducing cost and time for monitoring and research development of kelp seeding material. (Re)training with new datasets or extension of our available datasets allows for extended use for high-resolution microscopy images of other developmental stages or other seaweed species. This provides a baseline for future development of image analysis pipelines with, to our knowledge, the first high-throughput phenotyping pipeline for macro algal microscopic bright field images.

5.2. High model effectiveness and post-processing

For the evaluation of instance segmentation part of our pipeline, we followed the evaluation procedure of the Common Objects in Context (COCO) (Lin et al., 2014), using the mask mAP50 (the average of AP at IoU 0.50) and mask mAP50-95 (the average of AP at IoU thresholds ranging from 0.50 to 0.95 in 0.05 increments). What constitutes a "good" mAP score depends on the dataset and desired application. Our model obtained a mAP50 of 0.85 and mAP50-95 of 0.68 for masks, which is extremely good in comparison with the large COCO dataset and other models for microscopy instance segmentation. YOLOv11instance segmentation models scored between 0.32 – 0.48 mask mAP50-95 (Jocher & Qiu, 2024) on the COCO test set. To our knowledge, our model is the first to perform instance segmentation on microscopy images of seaweeds which complicates performance comparison. However, there are instance segmentation models which were trained on microscopy images with microalgae or single cells. In general, these datasets are more uniform due to smaller differences between instances in size and shape but obtained similar mAP50 and mAP50-95 scores to our model. For example, microscopic histopathology images, where cells can occur at high density and tend to overlap, obtained mAP scores for masks of mAP50 0.89 and mAP50-95 of 0.56 (Kang et al., 2024). Microalgae instance segmentation models obtained mAP50-95 scores of 0.72 and 0.54 (Chong et al., 2026; Sun et al., 2025).

We introduced a new post-processing method that merges instances over image slices (adjacent images) to retain the high-resolution of microscopic images when analysing large images. Our approach ISOM merges masks within and of neighbouring image slices based on the IoU threshold of the masks within the overlapping area of the bounding boxes of two instances. ISOM achieved significantly higher mAP50-95 scores compared to the standard approaches non-maximal suppression (NMS), non-maximal merging (NMM), or no merging. This merging method can be applied to any type of images that are too large to be processed without down sampling.

5.3. Saving time and resources while maintaining accuracy of manual annotation

Using our instance segmentation, the time and resources, especially labour costs, needed for research and quality control of seeding material are significantly reduced compared to manual image analysis. This can lead to faster knowledge development of biological processes of seedlings as more data can be analysed within the same time, allowing for labour intensive experiments that were previously not possible. Our phenotyping pipeline does not only save time but also reduces the amount of labour required from image acquisition to trait extraction. However, it is essential to retain data quality when replacing manual analysis with computer-based models. Therefore, for the trait extraction part of the pipeline, we compared the performance of our best model with manual annotation of three experts. The model analysed the images 18x faster and performed better than the experts. Our model detected more sporophytes, while the experts missed smaller sporophytes. Overall, the extracted trait values from our model matched ground-

truth measurements better. Additionally, inter and intra-personal differences can increase variability, which is reduced when using instance segmentation models.

High-throughput phenotyping of plant traits has become an increasingly established practice at different scales, such as field trials, greenhouses, and climate chambers and from plant to organ level (Araus & Cairns, 2014; Singh et al., 2016; W. Yang et al., 2020). The impact of these developments is expected to accelerate crop breeding (W. Yang et al., 2020). However, compared with agricultural crop research, such tools are not yet widely available for seaweed research and breeding. A recently developed phenotyping pipeline for macroalgae gametophytes demonstrates the potential of high-throughput phenotyping in seaweed but availability remains limited compared to terrestrial crops (Alves-Lima et al., 2026). We developed a high-throughput phenotyping pipeline based on microscopy images of early development of *S. latissima* sporophytes. The novelty of our method is not only its application to seaweed, but also high-throughput image acquisition and analysis at a microscopic scale. This provides a tool to assess large genotypic variety for early developmental traits enabling the generation of large dataset with objective and quantitative traits. For breeding this is valuable as throughput might otherwise limit the exploration of genetic variation. Our pipeline has potential broad application to phenotype different traits, such as (a)biotic stress resilience or early growth. For example, as sea temperatures are rising, selecting genotypes with improved heat tolerance would allow for more climate-proof biomass production. Previous research has already shown differences between genotypes for their temperature tolerance (Gonzalez et al., 2025; Nehr et al., 2025), highlighting the potential of high-throughput phenotyping.

5.4. Limitations

When assessing the performance of our model, we did not only consider quantitative assessment based on mAP-scores, but also qualitative assessment based on visualisations. We identified two characteristics that complicated instance segmentation, overlapping instances and the lack of clearly defined edges between sporophytes. Overlapping instances are a common problem in instance segmentation, such as in microscopy images of cervical cells (Jiang et al., 2023), steel surface damages (Bai et al., 2025), or as in our case overlapping sporophytes. For this reason, we choose YOLO11 because, compared to earlier versions of YOLO, enhanced feature pyramids and attention mechanisms improve the performance on partially occluded object (Khanam & Hussain, 2024). Post-processing approaches could further improve the model performance on overlapping instances. For example, Do-Net (Jiang et al., 2023) splits the detections into three layers to determine the separate objects or DA-Fusion (Park et al., 2025) combines RGB and depth data to separate occluding instances. Another important aspect is the input data. For example, when studying growth of sporophytes, overlap between individuals should be limited. By reducing the number of bouquets per coverslip or reducing the gametophyte size, leading to less sporophytes per bouquet will reduce the number of overlapping instances. The second characteristic that complicated instance segmentation was the lack of clearly defined edges between sporophytes. This can occur when the gametophyte is out of the focal plane, or when larger sporophytes are partially out-of-focus as blade tips start floating, while gametophytes remain attached to the substrate. This can be solved with destructive measurements by flattening the sample through the placement of a coverslip on top.

5.5. Future work and potential applications

In this study, we focused on models from the YOLO family. A promising direction for future research is the investigation of recently proposed architectures based on the Transformer paradigm, which have demonstrated strong performance across a range of vision tasks (Thisanke et al., 2023). In addition, semi-automatic segmentation approaches, where experts iteratively refine model-generated segmentation, could be integrated into the high-throughput phenotyping pipeline to improve segmentation accuracy and robustness while maintaining efficiency (Marinov et al., 2024). Finally, semi-supervised segmentation

methods, that leverage the use of both labelled and unlabelled data, can also be incorporated in the pipeline (Peláez-Vegas et al., 2026).

An important feature of our method is the availability of our datasets. This allows for expansion, specialisation and improvement of the models by other researchers. We propose several expansions of the dataset to broaden the application of our instance segmentation model. First, expansion of the dataset with colour images would allow for the usage of different methodologies for detection. Second, inclusion of more species of kelp would allow for broader use. As seedlings of commonly cultivated kelp species (*Alaria esculenta, Laminaria digitata, etc.*) are similar (Theodorou & Charrier, 2021) we expect high performance of our model on related species. However, including annotated images of different species would increase the robustness of the dataset. Lastly, the dataset contains less instances of earlier developmental stages, such as eggs and zygotes. Inclusion of eggs and zygotes into the dataset and pipeline would expand the use cases of this methodology for traits such as fertilization success, timing of gametogenesis and developmental success from egg to seedling.

Incorporation of this tool into experimental workflows increases the feasibility and economic viability of quantitative analysis of the microscopic seeding material. Research projects can assess e.g. the effect of responses to (a)biotic stresses such as light intensity, temperature, microbiome, exposure to flow, and attachment to substrates. Breeding companies may be able to make faster progress in phenotyping by screening larger genetic diversity. Other applications would be assessment of fertility rates, quality control of seeding material or comparing growth of seeding material to adult sporophyte size. We anticipate that this pipeline will function as the basis for high-throughput, accurate phenotyping for both commercial and research purposes.

References

Akyon, F. C., Onur Altinuc, S., & Temizel, A. (2022). Slicing Aided Hyper Inference and Fine-Tuning for Small Object Detection. *2022 IEEE International Conference on Image Processing (ICIP)*, 966–970. https://doi.org/10.1109/ICIP46576.2022.9897990

Alves-Lima, C., Barreto, L., Mónico, C., Gouvêa, L., Felix, F., Varga, B., Filipe, J., Camacho, R., Lymperaki, M., Alberto, F., & others. (2026). SAMMBA is a high-throughput pipeline for isolating and phenotyping macroalgal strains. *Algal Research*, 104547.

Araújo, R., Vázquez Calderón, F., Sánchez López, J., Azevedo, I. C., Bruhn, A., Fluch, S., Garcia Tasende, M., Ghaderiardakani, F., Ilmjärv, T., Laurans, M., Mac Monagail, M., Mangini, S., Peteiro, C., Rebours, C., Stefansson, T., & Ullmann, J. (2021). Current Status of the Algae Production Industry in Europe: An Emerging Sector of the Blue Bioeconomy. *Frontiers in Marine Science*, *7*, 626389. https://doi.org/10.3389/fmars.2020.626389

Araus, J. L., & Cairns, J. E. (2014). Field high-throughput phenotyping: The new crop breeding frontier. *Trends in Plant Science*, *19*(1), 52–61.

Bai, D., Yun, J., Jiang, D., Zhou, D., Gao, D., Li, G., & Ju, Z. (2025). HVA-YOLO: A real-time instance segmentation method under Hybrid Visual Attention applied to steel plate surface defect detection. *Mechanical Systems and Signal Processing*, *238*, 113191. https://doi.org/10.1016/j.ymssp.2025.113191

Bak, U. G., Mols-Mortensen, A., & Gregersen, O. (2018). Production method and cost of commercial-scale offshore cultivation of kelp in the Faroe Islands using multiple partial harvesting. *Algal Research*, *33*, 36–47. https://doi.org/10.1016/j.algal.2018.05.001

Bangare, S. L., Dubal, A., Bangare, P. S., & Patil, S. T. (2015). Reviewing Otsu's Method For Image Thresholding. *International Journal of Applied Engineering Research*, *10*(9), 21777–21783. https://doi.org/10.37622/IJAER/10.9.2015.21777-21783

Buschmann, A. H., Camus, C., Infante, J., Neori, A., Israel, Á., Hernández-González, M. C., Pereda, S. V., Gomez-Pinchetti, J. L., Golberg, A., Tadmor-Shalev, N., & Critchley, A. T. (2017). Seaweed production: Overview of the global state of exploitation, farming and emerging research activity. *European Journal of Phycology*, *52*(4), 391–406. https://doi.org/10.1080/09670262.2017.1365175

Cai, J., Lovatelli, A., Aguilar-Manjarrez, J., Cornish, L., Dabbadie, L., Desrochers, A., Diffey, S., Garrido Gamarro, E., Geehan, J., Hurtado, A., Lucente, D., Mair, G., Miao, W., Potin, P., Przybyla, C., Reantaso, M., Roubach, R., Tauati, M., & Yuan, X. (2021). Seaweeds and microalgae: An overview for unlocking their potential in global aquaculture development. In Food & A. O. of the U. Nations (FAO) (Eds), *FAO Fisheries and Aquaculture Circular* (Issue 1229, p. 48p.). https://doi.org/10.4060/cb5670en

Charrier, B., Abreu, M. H., Araujo, R., Bruhn, A., Coates, J. C., De Clerck, O., Katsaros, C., Robaina, R. R., & Wichard, T. (2017). Furthering knowledge of seaweed growth and development to facilitate sustainable aquaculture. *New Phytologist*, *216*(4), 967–975. https://doi.org/10.1111/nph.14728

Chong, J. W. R., Khoo, K. S., Ting, H.-Y., Iwamoto, K., Ma, Z., & Show, P. L. (2026). Translating pixels into identification: Cutting-edge microalgae detection and instance segmentation by leveraging YOLO models. *Ecological Informatics*, *95*, 103752. https://doi.org/10.1016/j.ecoinf.2026.103752

Chopin, T., Buschmann, A. H., Halling, C., Troell, M., Kautsky, N., Neori, A., Kraemer, G. P., Zertuche-González, J. A., Yarish, C., & Neefus, C. (2001). Integrating seaweeds into marine aquaculture systems: A key toward sustainability. *Journal of Phycology*, *37*(6), 975–986. https://doi.org/10.1046/j.1529-8817.2001.01137.x

Duarte, C. M., Wu, J., Xiao, X., Bruhn, A., & Krause-Jensen, D. (2017). Can Seaweed Farming Play a Role in Climate Change Mitigation and Adaptation? *Frontiers in Marine Science*, *4*. https://doi.org/10.3389/fmars.2017.00100

Eswar, I., Majumder, A., Raja mohamed, S. H., Logesh, K., Shek alavudeen, S. M. I., Hariprakash, M., & Sundarraj, D. K. (2025). Seaweed responses to ocean acidification: Global impacts on growth, biochemical composition, and CO2 mitigation potential. *2025*, *2*(1). https://doi.org/10.1007/s44289-025-00060-x

European Commission. (2020). Farm to fork strategy: For a fair, healthy and environmentally-friendly food system. *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions*, *381*, 1–9.

Fang, Y., & Zhong, B. (2023). Cell segmentation in fluorescence microscopy images based on multi-scale histogram thresholding. *Mathematical Biosciences and Engineering*, *20*(9), 16259–16278. https://doi.org/10.3934/mbe.2023726

Felzenszwalb, P. F., Girshick, R. B., McAllester, D., & Ramanan, D. (2010). Object Detection with Discriminatively Trained Part-Based Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(9), 1627–1645. https://doi.org/10.1109/TPAMI.2009.167

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 580–587. https://doi.org/10.1109/CVPR.2014.81

Gonzalez, S. T., Bell, T. W., Aydlett, M., Bailey, D., Jones, A., Kerr, H., & Lindell, S. (2025). Predicting heat tolerance in sugar kelp juvenile sporophytes via gametophyte heat stress testing. *Journal of Applied Phycology*, *37*(2), 1201–1212.

Guillard, R. R. L., & Lorenzen, C. J. (1972). YELLOW-GREEN ALGAE WITH CHLOROPHYLLIDE *C*[1,2]. *Journal of Phycology*, *8*(1), 10–14. https://doi.org/10.1111/j.1529-8817.1972.tb03995.x

Holdt, S. L., & Kraan, S. (2011). Bioactive compounds in seaweed: Functional food applications and legislation. *Journal of Applied Phycology*, *23*(3), 543–597. https://doi.org/10.1007/s10811-010-9632-5

Huang, M., Robbins, K. R., Li, Y., Umanzor, S., Marty-Rivera, M., Bailey, D., Yarish, C., Lindell, S., & Jannink, J. L. (2022). Simulation of sugar kelp (Saccharina latissima) breeding guided by practices to accelerate genetic gains. *G3: Genes, Genomes, Genetics*, *12*(3). https://doi.org/10.1093/g3journal/jkac003

Jiang, H., Zhang, R., Zhou, Y., Wang, Y., & Chen, H. (2023). DoNet: Deep De-Overlapping Network for Cytology Instance Segmentation. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 15641–15650. https://doi.org/10.1109/CVPR52729.2023.01501

Jocher, G., & Qiu, J. (2024). *Ultralytics YOLO11* (Version 11.0.0) [Computer software]. https://github.com/ultralytics/ultralytics

Kang, M., Ting, C.-M., Ting, F. F., & Phan, R. C.-W. (2024). ASF-YOLO: A novel YOLO model with attentional scale sequence fusion for cell instance segmentation. *Image and Vision Computing*, *147*, 105057. https://doi.org/10.1016/j.imavis.2024.105057

Khanam, R., & Hussain, M. (2024). *YOLOv11: An Overview of the Key Architectural Enhancements* (arXiv:2410.17725). arXiv. https://doi.org/10.48550/arXiv.2410.17725

Kraan, S. (2013). Mass-cultivation of carbohydrate rich macroalgae, a possible solution for sustainable biofuel production. *Mitigation and Adaptation Strategies for Global Change*, *18*(1), 27–46. https://doi.org/10.1007/s11027-010-9275-5

Lei, L., Yang, Q., Yang, L., Shen, T., Wang, R., & Fu, C. (2024). Deep learning implementation of image segmentation in agricultural applications: A comprehensive review. *Artificial Intelligence Review*, *57*(6), 149. https://doi.org/10.1007/s10462-024-10775-6

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds), *Computer Vision – ECCV 2014* (pp. 740–755). Springer International Publishing. https://doi.org/10.1007/978-3-319-10602-1_48

Luo, Z., Yang, W., Yuan, Y., Gou, R., & Li, X. (2024). Semantic segmentation of agricultural images: A survey. *Information Processing in Agriculture*, *11*(2), 172–186. https://doi.org/10.1016/j.inpa.2023.02.001

Marinov, Z., Jäger, P. F., Egger, J., Kleesiek, J., & Stiefelhagen, R. (2024). Deep Interactive Segmentation of Medical Images: A Systematic Review and Taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *46*(12), 10998–11018. https://doi.org/10.1109/TPAMI.2024.3452629

Mohan, D. S., & Chandar, K. P. (2025). A Comprehensive Review on Various Image Segmentation Methods with Applications. *2025 3rd International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, 1618–1623. https://doi.org/10.1109/ICSCDS65426.2025.11167803

Nehr, Z., Avia, K., Godfroy, O., Jacquemin, B., Mauger, S., Farhat, P., Gouhier, E., Coudret, J., Fort, A., Sulpice, R., & others. (2025). Genetic and transcriptomic analysis of a Saccharina latissima segregating F2 family detects groups of individuals with different responses to temperature stress. *Algal Research*, 104203.

Nichele, L., Persichetti, V., Lucidi, M., & Cincotti, G. (2020). Quantitative evaluation of ImageJ thresholding algorithms for microbial cell counting. *OSA Continuum*, *3*(6), 1417. https://doi.org/10.1364/OSAC.393971

Park, Y., Yoon, H. J., Kim, J., & Zhang, B. T. (2025). DA-Fusion: 2025 IEEE International Conference on Robotics and Automation, ICRA 2025. *2025 IEEE International Conference on Robotics and Automation, ICRA 2025, Proceedings - IEEE International Conference on Robotics and Automation*, 7490–7496. https://doi.org/10.1109/ICRA55743.2025.11128151

Peláez-Vegas, A., Aguilera-Martos, I., Mesejo, P., & Luengo, J. (2026). A Survey on Deep Learning-Based Semi-Supervised Semantic Segmentation. *ACM Computing Surveys*, *58*(12), 308:1-308:36. https://doi.org/10.1145/3806766

Polat, S., Trif, M., Rusu, A., Šimat, V., Čagalj, M., Alak, G., Meral, R., Özogul, Y., Polat, A., & Özogul, F. (2023). Recent advances in industrial applications of seaweeds. *Critical Reviews in Food Science and Nutrition*, *63*(21), 4979–5008. https://doi.org/10.1080/10408398.2021.2010646

Rebours, C., Marinho-Soriano, E., Zertuche-González, J. A., Hayashi, L., Vásquez, J. A., Kradolfer, P., Soriano, G., Ugarte, R., Abreu, M. H., Bay-Larsen, I., Hovelsrud, G., Rødven, R., & Robledo, D. (2014). Seaweeds: An opportunity for wealth and sustainable livelihood for coastal communities. *Journal of Applied Phycology*, *26*(5), 1939–1951. https://doi.org/10.1007/s10811-014-0304-8

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). *You Only Look Once: Unified, Real-Time Object Detection*. 779–788. https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Redmon_You_Only_Look_CVPR_2016_paper.html

Rønhovde, M. B., Monteiro, C., Dolan, D., Sjøtun, K., & Dahle, H. (2025). Northern Norway Sporophytes of *Saccharina latissima* Display Distinct Gene Expression Profiles in Response to Temperature and Photoperiod. *Ecology and Evolution*, *15*(5), e71455. https://doi.org/10.1002/ece3.71455

Singh, A., Ganapathysubramanian, B., Singh, A. K., & Sarkar, S. (2016). Machine learning for high-throughput stress phenotyping in plants. *Trends in Plant Science*, *21*(2), 110–124.

Sun, M., Jiang, J., Yang, Z., Kong, S., Qi, J., Shang, J., Luo, S., Sun, W., Wang, T., Wang, Y., Wang, Q., Dai, T., Chen, T., Zhang, J., Zhang, X., He, Y., Fu, P., Guan, Q., Zhou, S., … Yan, H. (2025). *VisAlgae 2023: A Dataset and Challenge for Algae Detection in Microscopy Images* (arXiv:2505.20687). arXiv. https://doi.org/10.48550/arXiv.2505.20687

Theodorou, I., & Charrier, B. (2021). Brown Algae: Ectocarpus and Saccharina as experimental models for developmental biology. In A. Boutet & B. Schierwater, *Handbook of Marine Model Organisms in Experimental Biology* (1st edn, pp. 27–47). CRC Press. https://doi.org/10.1201/9781003217503-2

Theodorou, I., Godfroy, O., Boscq, S., Billoud, B., Dusabyinema, Y., & Charrier, B. (2026). Radical cell identity bifurcation in Saccharina embryos coincides with the expression of newly acquired genes. *bioRxiv*, 2026–05.

Theodorou, I., Opsahl-Sorteberg, H.-G., & Charrier, B. (2021). Preparation of Zygotes and Embryos of the Kelp Saccharina latissima for Cell Biology Approaches. *Bio-Protocol*. https://doi.org/10.21769/BioProtoc.4132

Thisanke, H., Deshan, C., Chamith, K., Seneviratne, S., Vidanaarachchi, R., & Herath, D. (2023). Semantic segmentation using Vision Transformers: A survey. *Engineering Applications of Artificial Intelligence*, *126*, 106669. https://doi.org/10.1016/j.engappai.2023.106669

Umanzor, S., Li, Y., & Yarish, C. (2020). Effect of direct "seeding" binders and embryonic sporophyte sizes on the development of the sugar kelp, Saccharina latissima. *Journal of Applied Phycology*, *32*(6), 4137–4143. https://doi.org/10.1007/s10811-020-02277-z

Van Der Molen, J., Ruardij, P., Mooney, K., Kerrison, P., O'Connor, N. E., Gorman, E., Timmermans, K., Wright, S., Kelly, M., Hughes, A. D., & Capuzzo, E. (2018). Modelling potential production of macroalgae farms in UK and Dutch coastal waters. *Biogeosciences*, *15*(4), 1123–1147. https://doi.org/10.5194/bg-15-1123-2018

Vazquex Calderon, F., Sanchez Lopez, J., Guillen Garcia, J., & Avraamides, M. (2022). *An overview of the algae industry in Europe* (Scientific Analysis or Review KJ-07-22-673-EN-N (online)). Publications Office of the European Union. https://doi.org/10.2760/813113%20(online)

Veenhof, R. J., Grisenthwaite, R., Mair, A., Mitchell, E., Stanley, M. S., & Kumari, P. (2025). Optimizing nursery conditions of the commercial kelp Alaria esculenta. *Journal of Applied Phycology*, *37*(5), 3031–3042. https://doi.org/10.1007/s10811-025-03665-z

Visch, W., Rad-Menéndez, C., Nylund, G. M., Pavia, H., Ryan, M. J., & Day, J. (2019). Underpinning the Development of Seaweed Biotechnology: Cryopreservation of Brown Algae (*Saccharina*

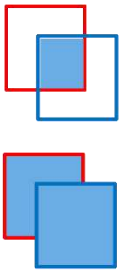
latissima) Gametophytes. *Biopreservation and Biobanking*, *17*(5), 378–386. https://doi.org/10.1089/bio.2018.0147

Yang, W., Feng, H., Zhang, X., Zhang, J., Doonan, J. H., Batchelor, W. D., Xiong, L., & Yan, J. (2020). Crop phenomics and high-throughput phenotyping: Past decades, current challenges, and future perspectives. *Molecular Plant*, *13*(2), 187–214.

Yang, X., Beyenal, H., Harkin, G., & Lewandowski, Z. (2001). Evaluation of biofilm image thresholding methods. *Water Research*, *35*(5), 1149–1158. https://doi.org/10.1016/S0043-1354(00)00361-4

Supplementary Information

$$\mathrm{IoU}_{\mathrm{box}} = \frac{\text{Area of overlap}}{\text{Area of union}}$$



$$\mathrm{IoU}_{\mathrm{mask}} = \frac{\text{Area of overlap}}{\text{Area of union}}$$



Fig.S1 *Intersection over union (IoU) entails a ratio between the area of overlap of the prediction with the ground truth, divided by the area of the merge of the prediction and the ground truth*

Table S1 Model information and effectiveness scores. Scores include Precision (True Positives / (True Positives + False Positives)), Recall (Ture Positives / (True Positives + False Negatives)), mean Average Precision at IoU 50 (mAP50), mean Average Precision a 0.05 step increase from IoU 50 to 95 (mAP50-95) and F1(2 x (Precision x Recall) / (Precision + Recall)).

Boxes					Masks					Model information		
Precision	**Recall**	**mAP50**	**mAP50_95**	**F1**	**Precision**	**Recall**	**mAP50**	**mAP50_95**	**F1**	**Name**	**Size (kB)**	**Inference (sec)**
0.962	1.000	0.992	0.938	0.980	0.962	1.000	0.992	0.942	0.980	*Alignment*	6138	0.067
0.996	**0.981**	**0.995**	**0.942**	**0.989**	**0.996**	**0.981**	**0.995**	**0.943**	**0.989**	*Alignment*	20291	0.073
0.980	0.981	0.994	0.945	0.981	0.980	0.981	0.994	0.930	0.981	*Alignment*	44413	0.113
0.908	**0.827**	**0.891**	**0.702**	**0.866**	**0.907**	**0.811**	**0.872**	**0.649**	**0.856**	*ADS*	6251	0.094
0.884	0.819	0.882	0.701	0.850	0.848	0.847	0.873	0.642	0.847	*ADS*	20383	0.099
0.857	0.833	0.858	0.690	0.845	0.849	0.825	0.846	0.626	0.836	*ADS*	44419	0.126
0.906	0.809	0.886	0.687	0.855	0.893	0.791	0.858	0.628	0.839	*ADS*	54796	0.129
0.832	0.867	0.873	0.714	0.849	0.832	0.867	0.866	0.671	0.849	*DDS*	6159	0.104
0.741	0.886	0.855	0.711	0.807	0.733	0.877	0.856	0.682	0.799	*DDS*	20319	0.105
0.822	0.873	0.864	0.705	0.846	0.822	0.873	0.857	0.679	0.846	*DDS*	44384	0.123
0.785	0.863	0.833	0.668	0.822	0.780	0.858	0.812	0.629	0.817	*>64 cells*	6135	0.087
0.786	0.830	0.833	0.667	0.808	0.783	0.830	0.818	0.632	0.805	*>64 cells*	20363	0.093
0.808	**0.873**	**0.837**	**0.668**	**0.839**	**0.789**	**0.853**	**0.811**	**0.643**	**0.819**	*>64 cells*	44359	0.112
0.782	0.760	0.793	0.651	0.771	0.729	0.838	0.791	0.618	0.780	*>64 cells*	54814	0.121

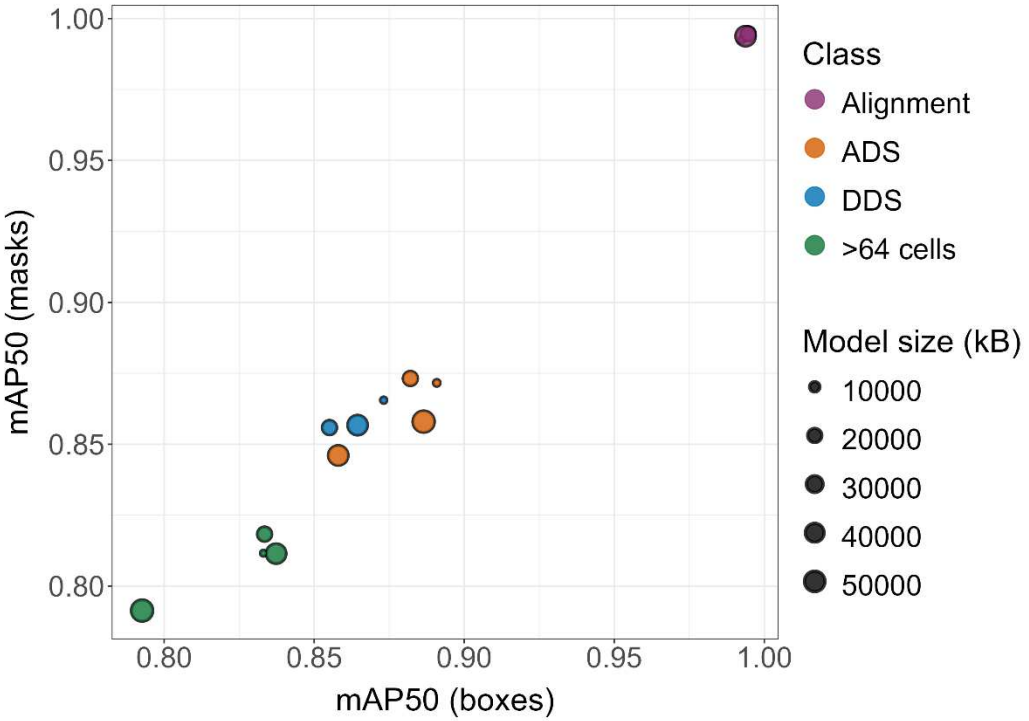


Fig.S2 *Mean Average Precision 50 (mAP50) for boxes and masks for four different model sizes and four different datasets (>64 cells, All developmental stages (ADS), Divided developmental stages (DDS), Alignment)*

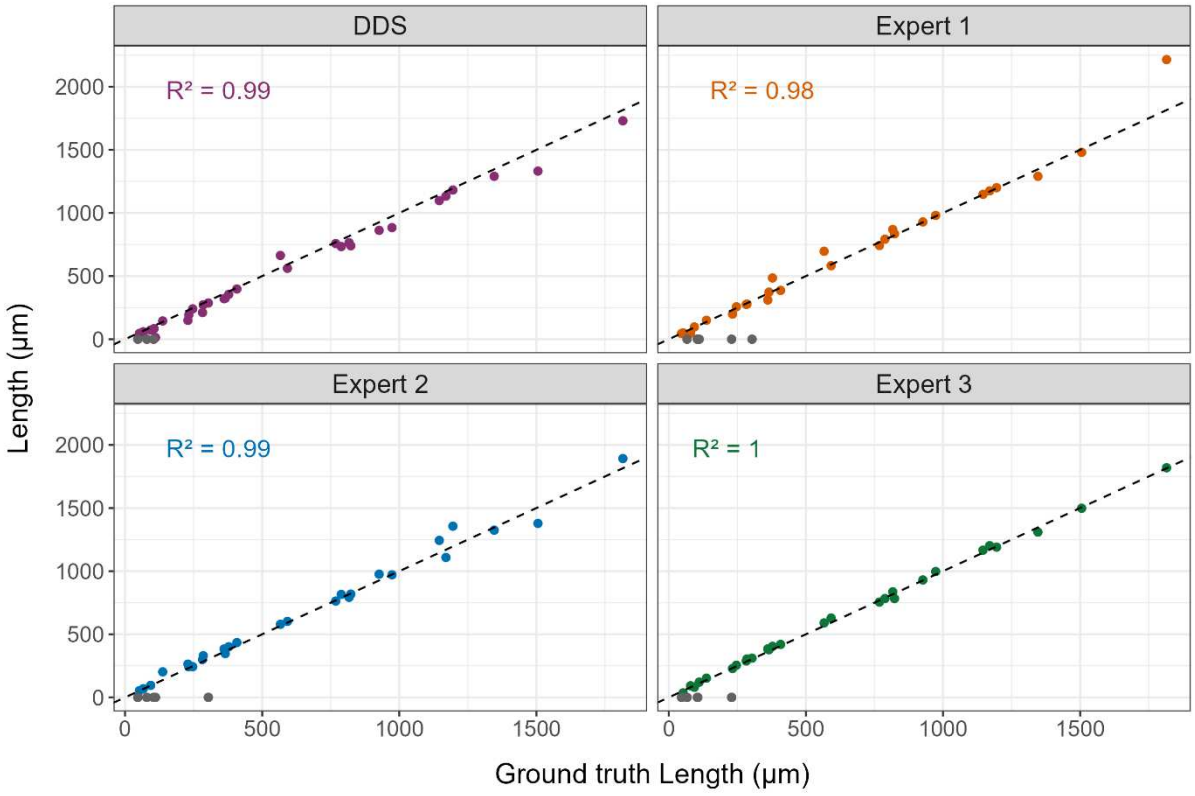


Fig.S3 *Comparison of the model (DDS) and three experts with the ground truth for length. Pearson correlations are indicated for the model and each expert. Grey points indicate sporophytes that were not detected by the model or expert and are not included in the calculation of the R^2*

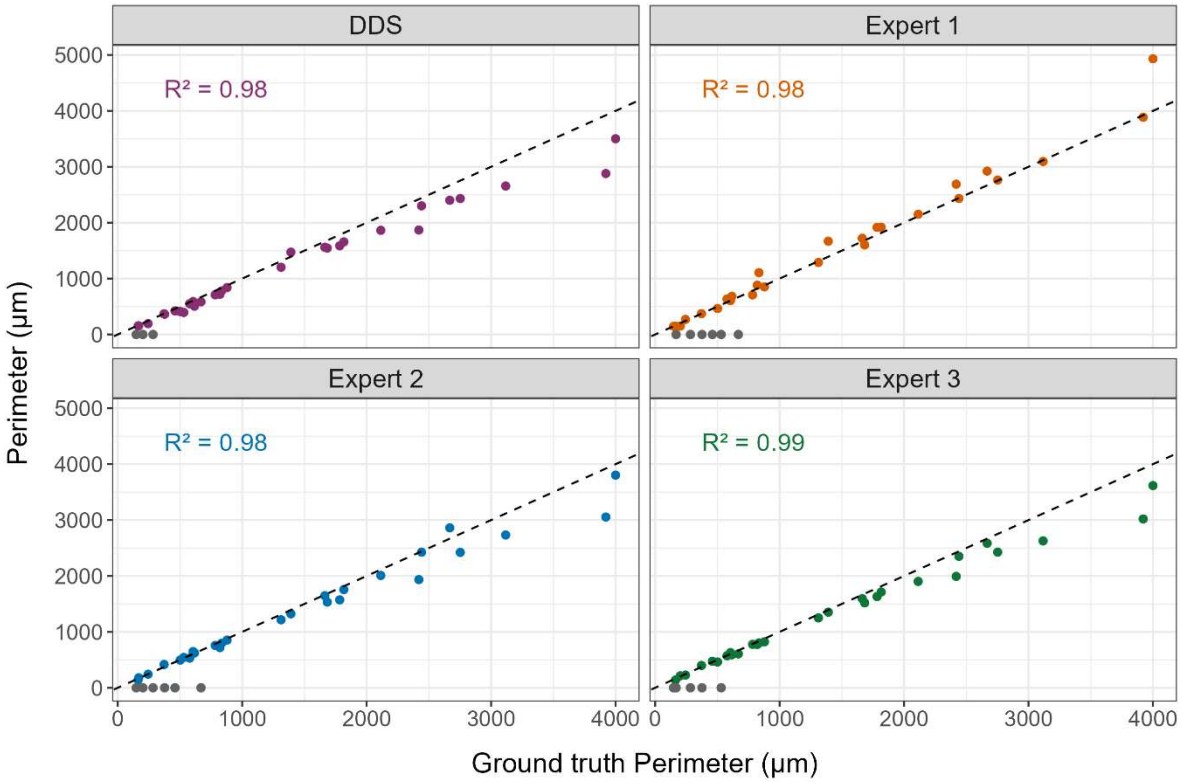


Fig.S4 *Comparison of the model (DDS) and three experts with the ground truth for perimeter. Pearson correlations are indicated for the model and each expert. Grey points indicate sporophytes that were not detected by the model or expert and are not included in the calculation of the R^2*

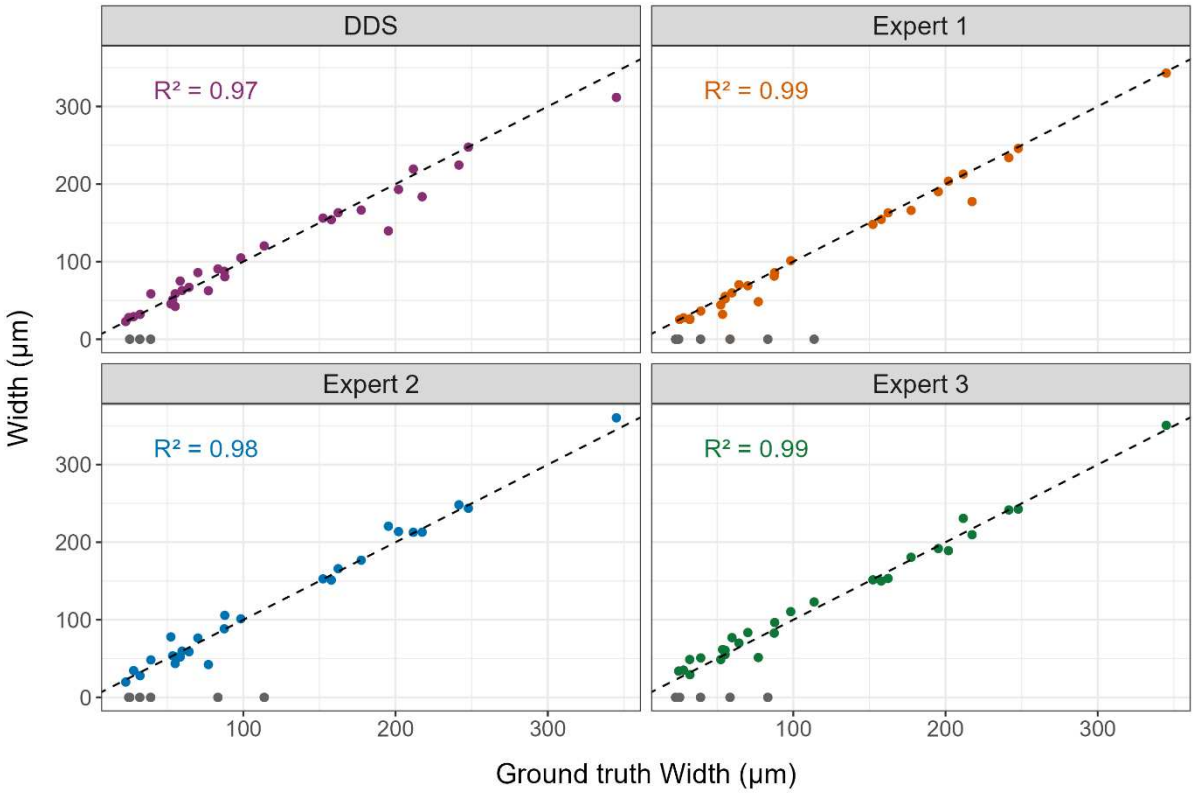


Fig.S5 *Comparison of the model (DDS) and three experts with the ground truth for width. Pearson correlations are indicated for the model and each expert. Grey points indicate sporophytes that were not detected by the model or expert and are not included in the calculation of the R^2*

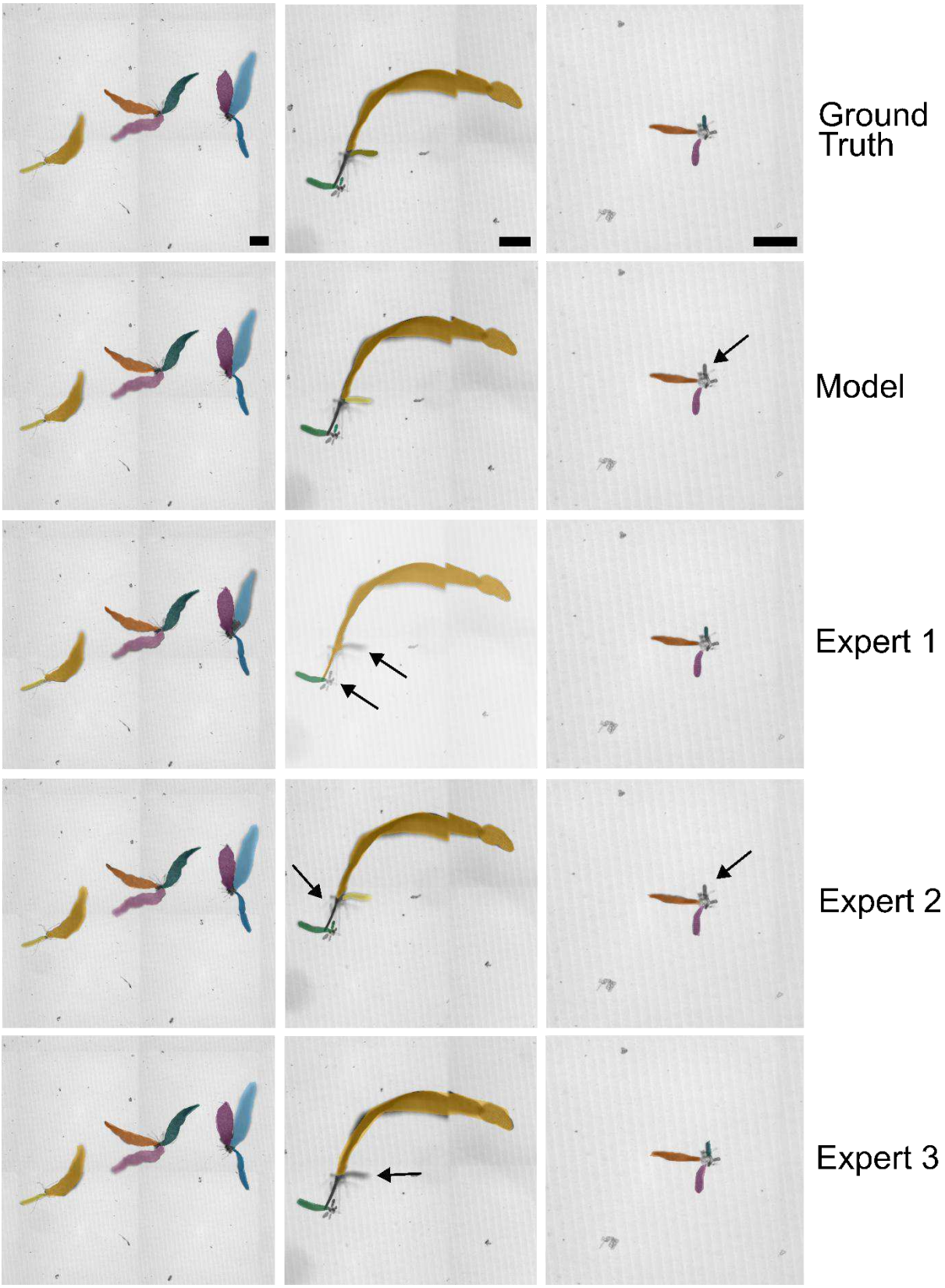


Fig.S6 *Examples of difference between model and expert annotations compared to the ground truth. Sporophytes that are missed by the experts are generally small, floating or lack distinct edges with larger sporophytes. The model predictions also miss the smaller, less distinct sporophytes. Arrows indicate missed annotations. Scale bar represents 200 µm*