# scientific **data**

Check for updates

OPEN

DATA DESCRIPTOR

# A Mycelium Dataset with Edge-Precise Annotation for Semantic Segmentation

Qianguang Yuan [1,2], Weizhen Liu [1,2 ✉], Yunfei Liu[3,4], Pin Li[5], Yuxuan Liu[5], Xiaohui Yuan[3,6], Nanqing Dong[7,8], Shengwu Xiong[9 ✉] & Yongping Fu[5,6 ✉]

With the increasing application of computer vision in mycology research, precisely segmenting mycelium and its edges in petri dish images remains a critical and underexplored task. This technology, accurately delineating mycelium boundaries, enables quantification of growth patterns, playing a crucial role in exploration of strain-related features, environmental adaptability, and physiological stimuli responses. The field confronts two bottlenecks, restricting real-world computer vision application. First, scarce public datasets impede development of mycelium-specific algorithms. Second, low contrast and high complexity of mycelium edges complicate annotation and segmentation processes. To address these bottlenecks, we established MyceliumSeg, the first large-scale benchmark dataset. MyceliumSeg contains: (i) 20,176 high-quality diverse images covering full growth cycle of four fungal species across multiple culture conditions; (ii) 567 pixel-level labeled samples generated with 37 person-days' manual effort through a mycelium annotation framework, including a multi-blind refined annotation guideline and a novel disagreement solution; (iii) a benchmark evaluating mainstream deep learning models under classic and boundary-aware segmentation metrics. MyceliumSeg serves as valuable resource for research on both mycology and segmentation algorithm.

## Background & Summary

Mycelium semantic segmentation represents a transformative approach in fungal research, offering unprecedented capabilities for large-scale analysis of hyphal network architectures. By accurately delineating mycelium boundaries, this technology enables the quantification of growth patterns that were previously unassessable by conventional methods[1,2]. The ability to capture subtle morphological variations empowers researchers to investigate strain-specific characteristics, monitor environmental adaptation processes, and evaluate physiological responses to various stimuli[3–5]. These advancements are driving innovation across disciplines, from ecological studies of fungal communities to the development of fungal-based biotechnological applications in medicine and agriculture. Moreover, the precision of semantic segmentation is particularly valuable for establishing correlations between morphological features and functional traits[6], thereby deepening our understanding of fungal biology and its practical applications.

However, the field currently faces two major bottlenecks. First, the lack of public benchmark datasets has created a critical resource gap, slowing down the development of segmentation algorithm and making it extremely difficult to reproduce methods or compare performance across research groups[7–11]. Second, the low contrast and high complexity of mycelium edges present dual challenges for both annotation and segmentation. Despite advances in deep learning for general image segmentation[12–14], the delicate and intertwined nature of hyphal edges and their blurred boundaries with the culture medium often cause traditional models to either under-segment, missing new hyphae, or over-segment, including unwanted noise. Manual annotation, which

[1]School of Computer Science and Artificial Intelligence, Wuhan University of Technology, Wuhan, 430070, China. [2]Sanya Science and Education Innovation Park, Wuhan University of Technology, Sanya, 572000, China. [3]Yazhouwan National Laboratory, Sanya, 572000, China. [4]School of Information Technology, Jilin Agricultural University, Changchun, 130118, China. [5]College of Mycology, Jilin Agricultural University, Changchun, 130118, China. [6]Jilin Provincial Key Laboratory of MycoPhenomics, Changchun, 130118, China. [7]Shanghai Innovation Institute, Shanghai, 200231, China. [8]Shanghai Artificial Intelligence Laboratory, Shanghai, 200232, China. [9]Interdisciplinary Artificial Intelligence Research Institute, Wuhan College, Wuhan, 430212, China. ✉e-mail: liuweizhen@whut.edu.cn; xiongsw@whut.edu.cn; fuyongping@jlau.edu.cn

| Species | Ganoderma lucidum | Pleurotus ostreatus | Trametes spp. | Ganoderma sinense |
|---|---|---|---|---|
| Culture medium | MYG | PDA | PDA | MYG |
| Temperature conditions | 15 °C, 25 °C | 25 °C | 25 °C | 25 °C |
| Storage conditions for sclerotium | Stored at 4 °C in wood-chip test tube | | | |
| Culture condition | 90-mm Petri dishes | | | |
| Light conditions | Dark incubation | | | |

**Table 1.** Summary of mycelium culture conditions. MYG stands for malt yeast glucose medium, and PDA stands for potato dextrose agar medium.

requires expert knowledge to distinguish hyphae from the background pixel by pixel, takes three to five times longer than labeling conventional biological images. This cycle of scarce data and edge-processing difficulties has limited the translation of algorithms from laboratory settings to real-world applications.

To address these challenges, we developed and released the first large-scale benchmark dataset for mycelium semantic segmentation, referred as to MyceliumSeg. It comprises 20,176 RGB mycelium images from four fungal species: *Ganoderma lucidum*, *Ganoderma sinense*, *Trametes* spp., and *Pleurotus ostreatus*. Images of mycelium were acquired under diverse culture conditions and span the full growth cycle from inoculation to full petri-dish coverage, capturing varied textures, colors, and morphological patterns. These images were captured using a self-developed and commercialized FPheno2000 imaging device[15], which employs a dual-light system: a 360° shadowless top light eliminates optical interference, while a bottom light enhances mycelial edge contrast and three-dimensional structure. This configuration overcomes the common issue of low edge contrast in traditional imaging, generating high-resolution images with clearer boundaries. The images accurately capture subtle morphological differences, such as the faint edges of newly grown hyphae, providing a solid foundation for pixel-level annotation and deep learning model training.

For data annotation, pixel-level annotations with fine edge labeling were provided for 567 representative samples covering the four fungal species. A multi-dimensional precise annotation framework was introduced to enhance annotation quality, featuring cross-expert labeling guidelines, conflict-detecting algorithms, and expert quality control teams to ensure high-quality, reproducible datasets. We tested three mainstream semantic segmentation algorithms including U-Net[16], DeepLabv3[17], and SegFormer[18] on this dataset. The results systematically revealed technical bottlenecks in hyphal edge segmentation: classic metrics such as mIoU and boundary-aware metrics like Boundary IoU[19], the 95th percentile of Hausdorff distance (HD95)[20,21], and Average Symmetric Surface Distance (ASSD)[22], highlighted the unique challenges of edge processing in fungal image analysis. This benchmark offers quantifiable ways to compare algorithm performance and identifies edge segmentation as a core challenge in fungal semantic analysis.

The dataset and benchmark system established in this study offer the first end-to-end solution for mycelium semantic segmentation, spanning data acquisition, fine-grained annotation, and algorithm evaluation. Their value lies not only in the scale of 20,176 images but also in the precise edge labeling that supports various algorithmic paradigms (fully supervised, semi-supervised, and self-supervised), particularly for edge-refined segmentation. In the future, this resource will facilitate applications such as automatically analyzing fungal phenotypes and monitoring mycelial states in real time during fermentation. It will also speed up the combination of deep learning and fungal research across different fields.

## Methods

In this section, we delve into the details of dataset construction and elaborate on the specific methods for data collection and the mycelium annotation. These methods are aimed at constructing a large-scale, high-quality mycelium dataset with pixel-level annotations and diverse data, so as to meet the research needs of precise segmentation.

**Data collection.** We collected 20,176 mycelium images with distinctive edge morphology. The samples spanning four fungal species, were stored at 4 °C in sawdust tubes, and were incubated in the dark on 90-mm Petri dishes with malt yeast glucose medium (MYG) or potato dextrose agar (PDA) culture medium in different temperatures (see Table 1). Image of these samples span diverse morphological characteristics, including growth stages, sclerotium colors, hyphal features (Fig. 1). The mycelium images were acquired using a mature, commercial data acquisition system named FPheno2000 developed by BORUIYUAN TECHNICAL (https://www.brytech.cn/). Following data acquisition process in Li *et al.*[15], we periodically placed mycelial petri dishes at a fixed position for image acquisition. Images with a resolution of 4,608 × 3,456 pixels are collected and saved in JPG format.

**Data annotation.** Due to inherently mycelium semi-transparent edges and low-contrast morphological features, precise pixel-level annotation and inter-annotator disagreements pose a significant concern. To achieve this, we proposed the mycelium annotation process comprising three steps: (a) a multi-blind refined annotation for manual error alleviation and pixel-level accuracy; (b) a disagreement disposal protocol containing a disagreement quantification method and disagreement solution; (c) expert review process ensuring the quality of the final annotation results (Fig. 2). Following this procedure, we produced 567 annotations, requiring a total of 37 person-days of manual effort. Representative annotation results are illustrated in Fig. 3.
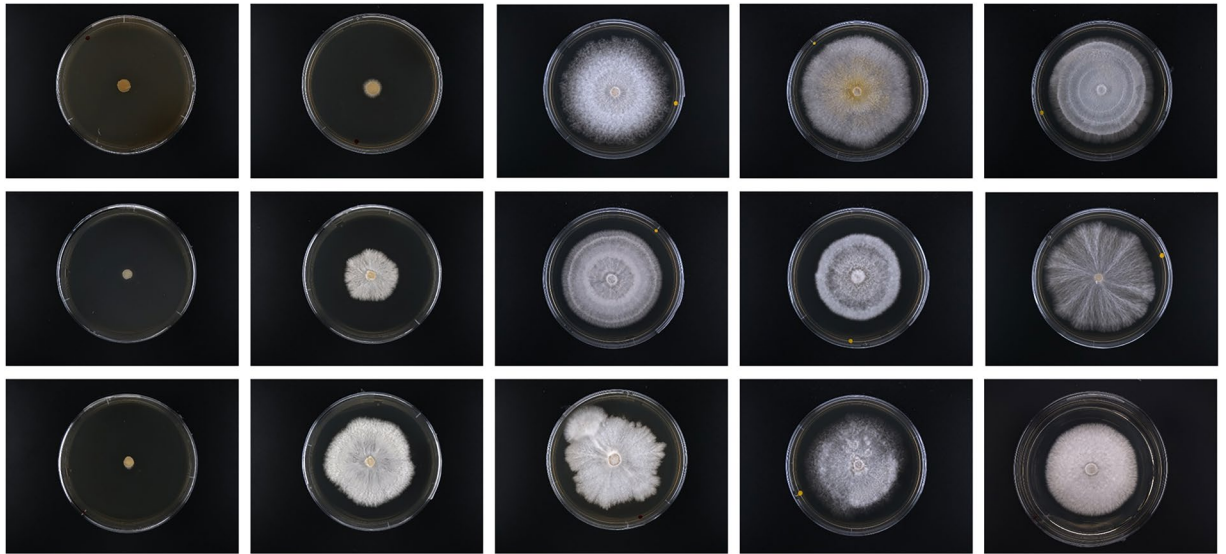
**Fig. 1** Visualization of morphological diversity within MyceliumSeg. Column 1 shows variations in sclerotium color during the activation and germination stage. Column 2 presents representative morphologies from the three subsequent growth stages (hyphal expansion, network building, and maturation). Columns 3–5 highlight diverse visual characteristics observed in the network building and maturation stages.

*Multi-blind refined annotation.* In the annotation process, multiple annotators independently label the same image without seeing other annotators' work. Specifically, only the mycelium growing around the sclerotium is considered as foreground, and the outermost fine edge of the mycelium is defined as the boundary of ground truth mask. Internal structural details or void regions of the mycelium are ignored. Other regions, including Petri dishes and culture medium are uniformly treated as background. Multi-blind annotation is employed to alleviate impact of potential visual confusion and blind spots caused by mycelium's weak features in single-annotator settings. In addition, a refinement operation, dedicated to label edge details after outlining the entire hyphal contour, is integrated into annotation process.

*Disagreement disposal protocol.* Disagreement is inevitable in the mycelium annotation with multiple annotators. We adopted a protocol combining disagreement quantification method and disagreement solution strategy. The disagreement quantification method comprises two parts. The first part is Mutual Average Symmetric Surface Distance (mASSD). mASSD quantifies the disagreement between a sample's designated annotation and all other annotations of that sample. The second part is sample level disagreement, which is defined as the sum of the mASSD values across all annotations of the same sample, and this total serves as an indicator of that sample's annotation difficulty.

The metric mASSD is based on ASSD, which is used to measures the average bidirectional distance between two contours. ASSD is calculated by sampling points along one contour, finding the nearest Euclidean distance from each point to the other contour, and averaging all distances[22]. An increased ASSD value between two contours signifies a correspondingly greater spatial divergence between them. As shown in Eq. 1, $ASSD_{(i,j),k}$ represents the ASSD between annotators i and j on sample k:

$$ASSD_{(i,j),k} = ASSD(S_{i,k}, S_{j,k}) = \frac{1}{|S_{i,k}| + |S_{j,k}|}\left(\sum_{y \in S_{i,k}} min_{x \in S_{j,k}} \|x - y\| + \sum_{y \in S_{j,k}} min_{x \in S_{i,k}} \|x - y\|\right),$$ (1)

where $S_{i,k}$ denotes point set of the k-th sample's contour from annotator *i*, and point in the set is represented by *x* and *y*. Following all pairwise ASSDs have been obtained, the mASSD for a designated annotation is defined as the mean of its ASSD values to every other annotation of the same sample. $mASSD_{j,k}$ quantifies the average disagreement between the annotation of sample *k* produced by annotator *j* and the annotations of the same sample produced by all other annotators, *i.e.*

$$mASSD_{j,k} = \frac{1}{|N - 1|}\sum_{i \neq j} ASSD_{(i,j),k}, \ (i = 1, 2, \ldots, N),$$ (2)

where *i* is the sequence of annotator and *N* is the total annotators. Sample level disagreement, indicating annotation difficulty of a sample, is calculated as the sum of that sample's mASSD values across all annotators (Eq. 3).

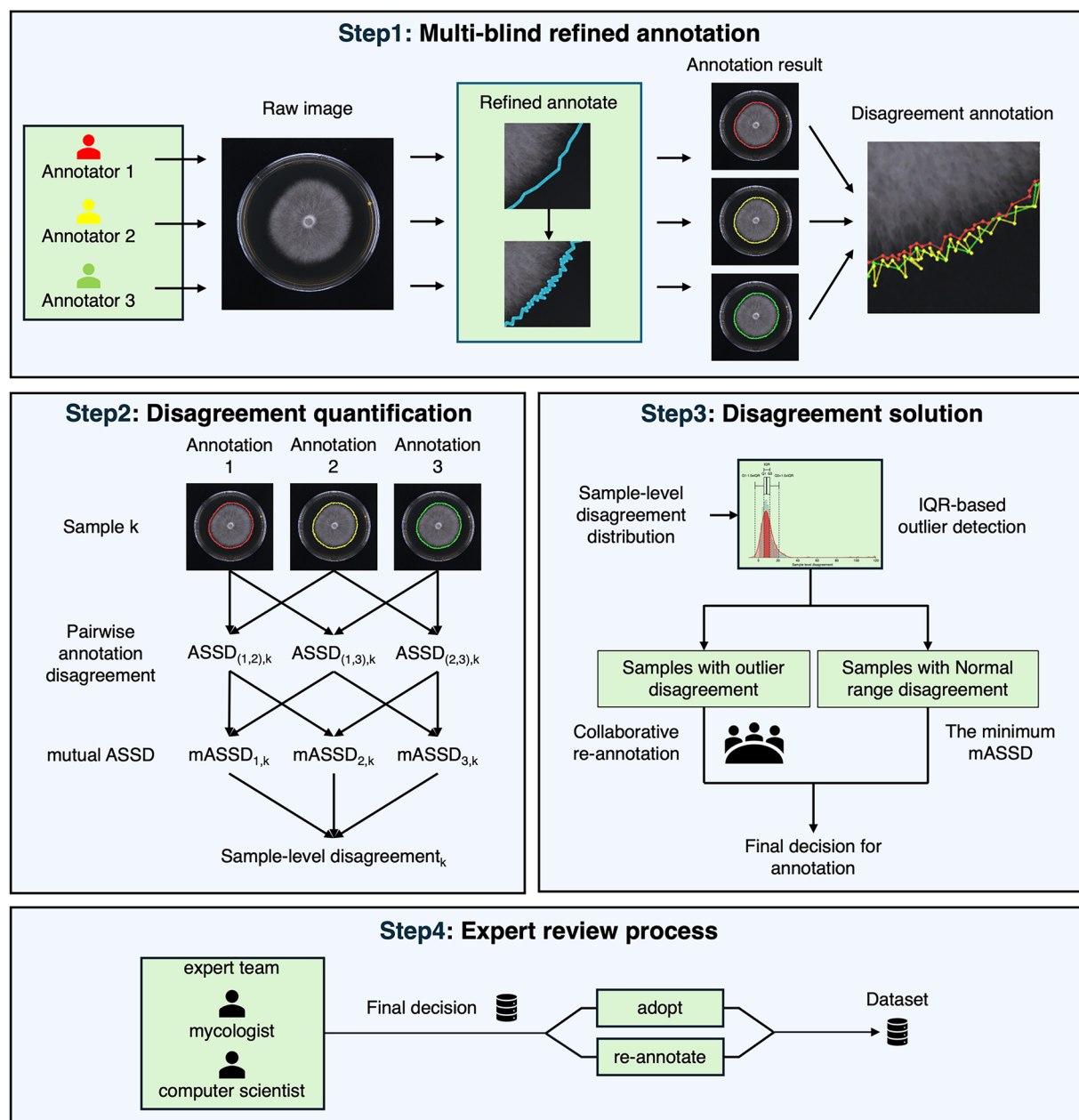$$Sample\ level\ Disagreement_k = \sum_j mASSD_{j,k}, \ (j = 1, 2, \ldots, N)$$ (3)

**Fig. 2** Annotation workflow. (**a**) Multi-blind refined annotation. Each image is first labelled independently by multiple annotators who cannot see one another's work. They draw a coarse contour of each mycelium sample and then refine the boundary pixel by pixel. (**b**) Disagreement disposal protocol consists of disagreement quantification method and solution. Pixel-level mismatches among multiple refined annotations are quantified. Different disagreement-handling solutions are applied to each sample based on the quantified disagreement results. (**c**) Expert review process is used to ensure the annotation quality.

After the disagreement quantification method, we assembled a collaborative panel combining with statistical analyses to resolve the disagreements. An interquartile range (IQR)–based outlier detection was applied to the distribution of sample level disagreement values to identify samples exhibiting elevated annotation discrepancies[23]. The panel would review the annotations of these samples to determine the necessity of re-annotation and would re-annotate together to ensure objective and accurate results. Moreover, for samples with disagreement scores in the normal range, the annotation with the lowest mASSD is chosen as the final annotation. This approach ensure that final annotation diverges minimally from all other annotations.

*Expert review process.* An expert team comprising mycologist and computer scientist reviewed and approved the annotations. If discrepancies or ambiguities remained, they would collaboratively re-annotate the data to ensure that these valuable cases are annotated with high precision and fully utilized.
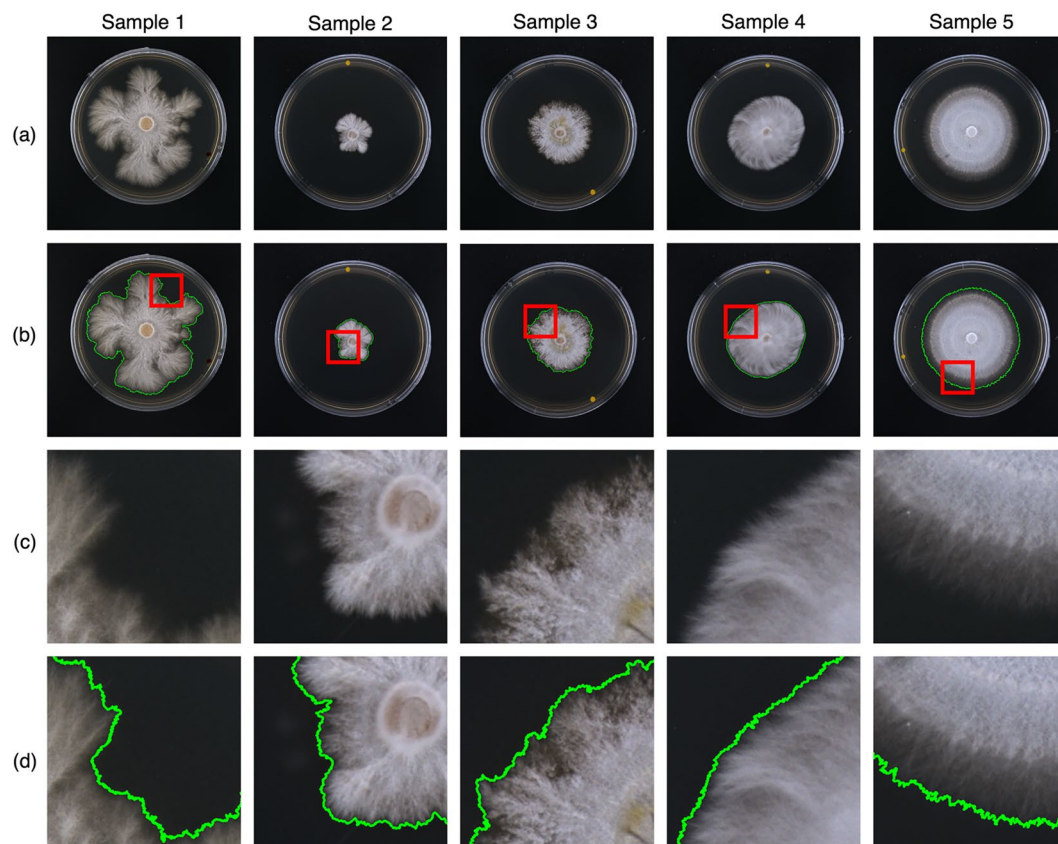
**Fig. 3** Overview of the raw image and its annotation at global and local scales. (**a**) Original image. (**b**) Edge map of the full annotation on the original image with the magnified region indicated. (**c**) Magnified view of the local original image; (**d**) Corresponding magnified view of the annotation edge.

## Data Records

The dataset is accessible for download at Zenodo[24]. MyceliumSeg comprises five parts: 'labeled-GL', 'labeled-GS_PO_TS', 'labeled-MYG_PDA_TEMP', 'unlabeled-GL', and 'unlabeled-GS_PO_TS'. The 'labeled-GL' folder comprises 507 labeled *Ganoderma lucidum* images, which is divided into two subfolders, 457 images for 'trainset' and 50 images for 'testset'. The 'labeled-GS_PO_TS' folder comprises 30 labeled images of *Ganoderma sinense*, *Trametes* spp., and *Pleurotus ostreatus*. The image is equally divided into three subfolders: 'GS', 'TS', and 'PO'. The 'labeled-MYG_PDA_TEMP' folder comprises 30 labeled images, equally split (10 each) among MYG-based medium (MYG), PDA-based medium (PDA), and 15 °C incubation (TEMP15), and is organized into the 'MYG', 'PDA', and 'TEMP15' subfolders. Each of these labeled subfolders further comprises an 'image' and a 'mask' folder: the 'image' folder stores raw images in '.jpg' format, whereas the 'mask' folder holds the pixel-wise annotations in binary '.png' files (0 for background, 1 for mycelium). Filenames are identical across the paired image and mask files. The 'trainset' contains files numbered from '00000001' to '00000457'. The 'testset' contains files numbered from '00000458' to '00000507'. The 'GS', 'PO', 'TS', 'MYG', 'PDA' and 'TEMP15' contains files numbered from '00018428' to '00018437', '00018438' to '00018447', '00018448' to '00018457', '00018458' to '00018467', '00018468' to '00018477' and '00018478' to '00018487', separately. The unlabeled data part consists of 'unlabeled-GL' and 'unlabeled-GS_PO_TS'. The former part consists of seven subfolders, 'unlabeled-GL1' through 'unlabeled-GL7', which hold 17,920 *Ganoderma lucidum* original unlabeled '.jpg' images with sequential filenames ranging from '00000508' to '00018427'. The latter part 'unlabeled-GS_PO_TS' contains 1689 unlabeled images of *Ganoderma sinense*, *Trametes* spp., and *Pleurotus ostreatus*, with filenames consecutively numbered from '00018488' to '00020176'.

## Technical Validation

This section presents statistical analysis of the collected data from lifecycle, sclerotium and hyphal visual features. The disagreement distribution is demonstrated with boxplots. What's more, the dataset is benchmarked across several seminal deep learning-based segmentation architectures.

**Data statistical analysis.** MyceliumSeg provides image data that comprehensively span all stages of mycelial growth, showcasing the unique morphological diversity characteristic of each phase (Fig. 1).

*Lifecycle analysis.* Table 2 provides statistics on the mycelial growth stages. Since the images were acquired throughout the mycelial cultivation process, the relative frequency of data in each growth stage proportionally

| Morphological characteristic | Subclass | Counts |
|---|---|---|
| Growth stages | Sclerotium activation and germination stage | 2,204 (10.92%) |
| | Primary hyphal expansion stage | 2,652 (13.14%) |
| | Hyphal network construction stage | 10,894 (53.99%) |
| | Mycelial maturation transition stage | 4,426 (21.94%) |
| Total | | 20,176 |

**Table 2.** Distribution of mycelial growth stage frequencies in the dataset.

| Morphological characteristic | Subclass | Counts |
|---|---|---|
| Sclerotium colors | Yellow | 1,122 (50.91%) |
| | Blended | 342 (15.52%) |
| | Grey | 317 (14.38%) |
| | Brown | 303 (13.75%) |
| | Black | 120 (5.44%) |
| Total | | 2,204 |

**Table 3.** Distribution of sclerotium color frequencies in the dataset.

| Hyphal features | Description |
|---|---|
| Uniform density distribution | Biomass density remains even from center to edge. |
| Concentric density zonation | Alternating dense-sparse rings record periodic growth pulses. |
| Centripetal densification | Inner core compacts and darkens as mycelium growth. |
| Irregular edges | Edge shows lobes, waves, or serrations. |
| Peripheral densification | Thick, active growth band forms at outer rim. |
| Hyphal pigmentation | Pigment deposition produces yellow-to-dark brown coloring. |
| Heterogeneous density distribution | Patchy high- and low-density clusters across mycelium. |
| Wrinkling | Surface folds or ridges from uneven aerial-hypha shrinkage. |
| Rhizomorph | Rope-like cords carry nutrients and anchor mycelium. |
| Spiral stratification | Concentric bands follow outward helicoidal spiral. |
| Internal concavity | Central depression from autolysis and water loss. |

**Table 4.** Mycelial visual features and descriptions.

reflects the temporal duration of those phases. The majority of data (10,894, 53.99%) were acquired during hyphal network construction stage, followed by the next largest share (4,426, 21.94%) collected in mycelial maturation transition stage. Data from these two stages exhibit pronounced structural and color visual features. In contrast, the smallest subset of images (2,204, 10.92%) was obtained during sclerotium activation and germination stage, characterized primarily by color-based visual attributes. The remaining images (2,652, 13.14%) correspond to primary hyphal expansion stage, whose visual characteristics lack distinctive analytical significance[25–27].

*Sclerotium analysis.* In the sclerotium activation and germination stage (2,204, 10.92%), the visual features are reflected in sclerotial color. Overall, 66.43% of sclerotium appear yellow (see Table 3). Among these, 50.91% are pure yellow and 15.52% are a yellow and black blend. The remaining sclerotium are 14.38% gray, 13.75% brown, and 5.44% black. These images illustrate the diversity of sclerotium color patterns prior to hyphal growth.

*Hyphal feature analysis.* In hyphal network construction (10,894, 53.99%) and mycelial maturation transition stage (4,426, 21.94%), mycelium display distinctive structural or color signatures. Table 4 lists the visual features present in the dataset and describes them, while Table 5 summarizes their distribution across the 15,320 images. 7,295 images (47.62%) exhibit a uniform density distribution, whereas 3,248 images (21.20%) show concentric density zonation. Centripetal densification is evident in 1,283 images (8.37%), and peripheral densification in 557 images (3.64%). Edge morphology statistics reveal 1,169 mycelium (7.63%) with irregular edge. Less frequent yet informative traits include hyphal pigmentation (441), heterogeneous density distribution (400), wrinkling (297), rhizomorph (275), spiral stratification (238), and internal concavity (117), each accounting for under 3% of the dataset.

**Evaluation metric.** In the design of evaluation system, dual considerations were incorporated: first, accounting for the methodological significance of edge segmentation precision in mycelium segmentation research.

| Morphological characteristic | Subclass | Counts |
|---|---|---|
| Hyphal features | Uniform density distribution | 7,295 (47.62%) |
| | Concentric density zonation | 3,248 (21.20%) |
| | Centripetal densification | 1,283 (8.37%) |
| | Irregular edges | 1,169 (7.63%) |
| | Peripheral densification | 557 (3.64%) |
| | Hyphal pigmentation | 441 (2.88%) |
| | Heterogeneous density distribution | 400 (2.61%) |
| | Wrinkling | 297 (1.94%) |
| | Rhizomorph | 275 (1.80%) |
| | Spiral stratification | 238 (1.55%) |
| | Internal concavity | 117 (0.76%) |
| Total | | 15,320 |

**Table 5.** Distribution of hyphal characteristic frequency in the dataset.

Second, addressing the limitations of classical segmentation metrics, which exhibit heightened sensitivity to mask interior regions while demonstrating insufficient sensitivity to edge segmentation accuracy.

The classical segmentation metrics used to benchmark the model are the F1-score (Eq. 5) and Intersection-over-Union (IoU) (Eq. 8). Because the dataset can be foreground-sparse, we report these metrics for the foreground class by default, *i.e.*, the mycelium. F1-score of mycelium is defined as follows:

$$Mycelium \ F_1 = 2 \times \frac{Precision_f \times Recall_f}{Precision_f + Recall_f}, \tag{5}$$

where $Precision_f$ is the proportion of truly foreground pixels among all pixels predicted as foreground, and $Recall_f$ is the proportion of ground-truth foreground pixels that are correctly identified by the model. The $Precision_f$ and $Recall_f$ are defined as follows:

$$Precision_f = \frac{TP}{TP + FP}, \tag{6}$$

$$Recall_f = \frac{TP}{TP + FN}, \tag{7}$$

where true positive (TP), false positive (FP) and false negative (FN) are represent the number of foreground pixels predicted as foreground, background pixels predicted as foreground and foreground pixels predicted as background. With these quantities, IoU of mycelium is expressed in Eq. (8):

$$Mycelium \ IoU = \frac{TP}{TP + FP + FN} \tag{8}$$

To resolve the limitation of classic metrics, edge accuracy quantification metrics including Boundary IoU[19], HD95[20,21] and ASSD[22] (Eqs. 9–13) were systematically integrated to enable precise evaluation of edge segmentation performance from different aspects. Boundary IoU calculates the intersection-over-union for mask pixels within a certain distance from the corresponding ground truth or prediction boundary contours, i.e.

$$Boundary \ IoU(G, P) = \frac{|(G_d \cap G) \cap (P_d \cap P)|}{|(G_d \cap G) \cup (P_d \cap P)|}, \tag{9}$$

where G is ground truth binary mask, P is prediction binary mask, and boundary regions $G_d$ and $P_d$ are the sets of all pixels within d pixels distance from the ground truth and prediction contours respectively. Boundary dilation ratio is the hyper-parameter that specifies the proportion of $d$ relative to the image diagonal, and a smaller ratio imposes a stricter criterion on boundary segmentation. HD95 and ASSD are used to provide comprehensive evaluation for the results of edge segmentation from the view of the similarity between two masks. HD95 is used for measuring the impact of outliers or noise. It is defined as:

$$HD95(G, P) = max(HD95_{GP}, HD95_{PG}), \tag{10}$$

$$HD95_{GP} = percentile_{95}(\min_{b \in S(P)} \|a - b\|), \ \forall \ a \in S(G), \tag{11}$$

$$HD95_{PG} = percentile_{95}\left(\min_{a\in S(G)}\|b-a\|\right), \forall\, b \in S(P),$$

(12)

where S(·) represents the set of points on the surface of mask, ||·|| denotes the Euclidean distance between two points, and percentile$_{95}$ is the function returning the 95th percentile of distances. ASSD is a metric used to measure the average distance between the surfaces of ground truth and prediction masks, and it is mathematically formulated as:

$$ASSD(G, P) = \frac{1}{|S(G)| + |S(P)|}\left(\sum_{a\in S(G)}\min_{b\in S(P)}\|a-b\| + \sum_{b\in S(P)}\min_{a\in S(G)}\|b-a\|\right)$$

(13)

**Implementation details.** The main goals of the experimental design on MyceliumSeg dataset are two folds. First, we aim to evaluate the performance of representative segmentation baseline on the dataset for boundary-aware segmentation measurement. Second, we aim to evaluate the robustness of model in mycelium boundary-aware segmentation under different fungal species and culture conditions. By achieving these, we hope to establish a benchmark for future work and promote further research in this field.

To cover both CNN- and Transformer-based architectures, we benchmarked three representative segmentation baselines, U-Net[16], DeepLabv3[17], and SegFormer[18]. For a fair comparison, we used AdamW[28] ($\beta_1 = 0.9$, $\beta_2 = 0.999$) as the base optimizer with batch size of 4 per GPU for all models but allowed architecture-specific settings. We largely retained the default hyper-parameter settings in MMSegmentation[29]. For CNN-based architectures, vanilla U-Net and DeepLabv3 with ResNet-50 backbone were initiated with a learning rate of 2e-4 and a weight decay of 1e-5. The poly learning strategy with power of 0.9 was adopted. For Transformer-based architecture, SegFormer with MiT-B0 backbone adopted a lower initial learning rate of 6e-5, a higher weight decay of 1e-2, and a 3,000-iteration linear warm-up (warmup ratio = 10e-6) before switching to a polynomial schedule with power of 1.0. We trained all models for 50,000 iterations and report the last performance measured in mycelium IoU, mycelium F1-score, HD95, ASSD and Boundary IoU. Boundary dilation ratio of Boundary IoU was fixed at 0.001 to impose a more stringent criterion on edge segmentation. All experiments are implemented by PyTorch[30] based on MMSegmentation using four NVIDIA 4090 GPUs with 24 G memory.

The baseline models were constructed via fully supervised learning using 507 annotated images of *Ganoderma lucidum* (457 for training and 50 for testing). The model with the best performance was selected for multi-dimensional robustness evaluation. For the cross-species dimension, the model was directly applied to images of *Ganoderma sinense*, *Pleurotus ostreatus*, and *Trametes* spp. (10 images per species) for inference to assess its robustness, respectively. For the temperature dimension, model inference tests were conducted on 10 images of *Ganoderma lucidum* cultured at 15 °C and the performance of baseline was referred as the result of 25 °C. For the culture medium dimension, model inference tests were conducted on 10 images of *Ganoderma lucidum* grown on MYG plates and 10 images of *Trametes* spp. grown on PDA plates.

**Disagreement solution.** We analyzed the distributions of the disagreement-related metrics and presented them in box plots accordingly to assess the consistency of different annotators' results. Any instances with significant disagreement would be addressed to ensure annotation quality. The annotation disagreements among annotators, two computer science researchers and an externally contracted annotator, were quantified by analyzing the distributions of mASSD and sample level disagreement values.

The distributions of the disagreement-related metrics are presented in Fig. 4. In Fig. 4(a), the ASSD values between annotator 1 and annotator 2 are the lowest among all annotator pairs, indicating the highest level of agreement. In Fig. 4(b), annotator 1 achieves the lowest mASSD value, reflecting minimal relative disagreement with all other annotators. Additionally, guided by the sample level disagreement metric, a subset of high-disagreement, challenging samples was identified for collaborative annotation adjustments or re-annotating.

**Benchmark evaluation.** Table 6 presents the test results of three segmentation models, UNet, DeepLabv3, and SegFormer, after they underwent supervised training using the trainset comprising 457 annotated images. While all three algorithms demonstrate respectable performance in global segmentation metrics such as Mycelium F1-score and Mycelium IoU (all scores exceeding 84%), their performance in critical boundary-focused metrics, including Boundary IoU, HD95, and ASSD, was notably insufficient. Specifically, SegFormer achieves the highest score 28.60% of Boundary IoU, whereas U-Net and DeepLabv3 achieve 27.74% and 27.31%, respectively. The differences among the three models are minimal. The highest score indicates that SegFormer delivers finer edge segmentation than the other models, whereas the small margin suggests that the existing mainstream architectures remain inadequate for stringent fine-edge segmentation tasks. In contrast, DeepLabv3 outperforms U-Net and Segformer on the score of HD95 metric, achieving 63.53 compared with 139.34 and 75.95, respectively. The lowest HD95 score for DeepLabv3 indicates far less impact to complexity boundary outliers and local extreme noise, whereas the much higher score for U-Net reflects its limited ability to delineate fine boundaries under complexity edge features or noisy conditions. As for ASSD metric, SegFormer records 15.28, while U-Net and DeepLabv3 obtain 45.44 and 18.50, respectively. The lowest ASSD score indicates that SegFormer's predicted masks achieve the greatest similarity to the ground truth and that SegFormer is better able to capture the geometric characteristics of the mycelium.

The visualization in Fig. 5 qualitatively illustrates the quantitative trends reported in Table 6. In rows 1 and 2, where mycelium boundary is clear, all three models realize high Mycelium F1-score and IoU, and SegFormer
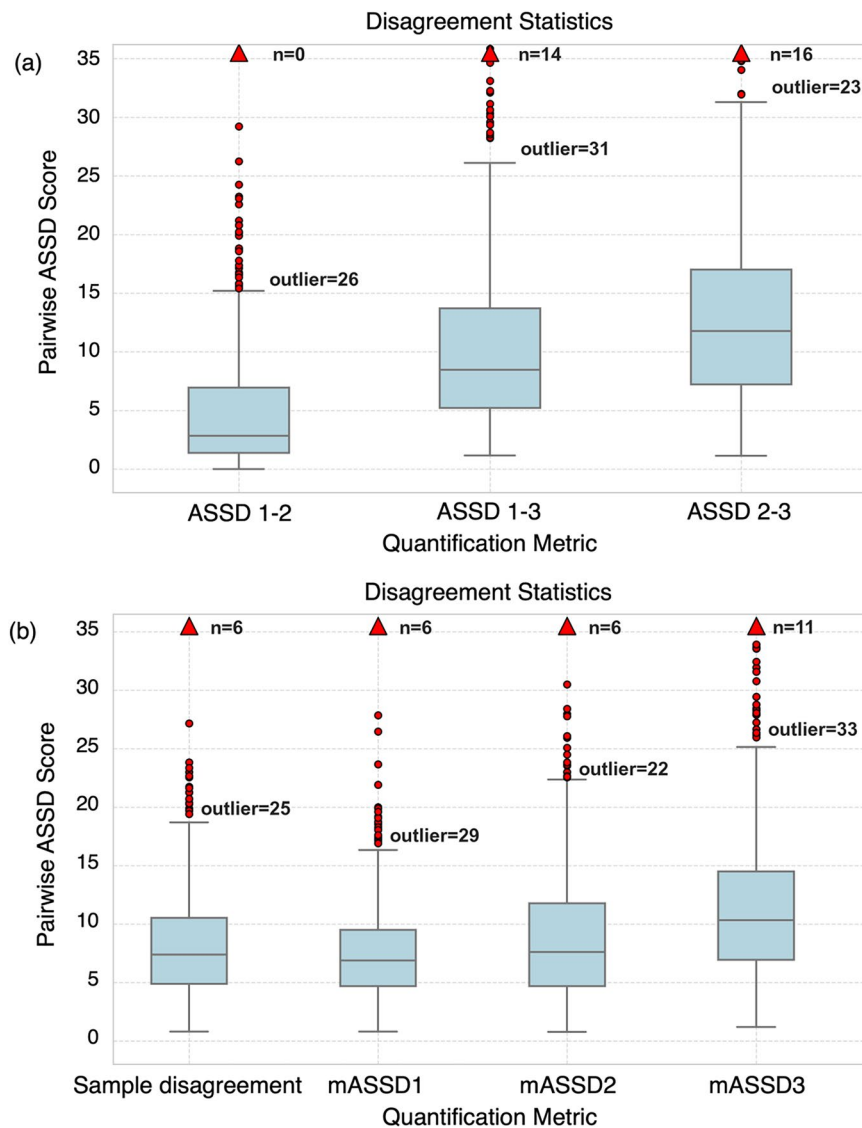
**Fig. 4** Distribution chart of disagreement related metrics. 1, 2, and 3 denote annotator indices corresponding respectively to computer researchers experienced in mycelium cultivation, computer researchers without cultivation experience, and outsourced personnel. (**a**) Distribution of ASSDs. (**b**) Distribution of mASSD and sample-level disagreements.

| Model | Mycelium F1-score ↑ (%) | Mycelium IoU ↑ (%) | Boundary IoU ↑ (%) | HD95 ↓ (px) | ASSD ↓ (px) |
|-------|------------------------|--------------------|--------------------|-------------|-------------|
| U-Net | 91.61 | 84.52 | 27.74 | 139.34 | 45.44 |
| DeepLabv3 | 97.33 | 94.80 | 27.31 | 63.53 | 18.50 |
| SegFormer | 97.55 | 95.22 | 28.60 | 75.95 | 15.28 |

**Table 6.** The performance of various mainstream models.

achieves the lowest ASSD. Nevertheless, the Boundary IoU values of the three remain tightly clustered near 28% without following the tendency of ASSD. It demonstrates that although the ability of capturing geometric characteristics has improved with successive architectural updates, precise edge alignment has not yet benefited from that. Row 3 describes a sample with jagged, low-contrast borders. The visible drift in the predictions reflects their elevated HD95 scores. SegFormer lowers the score compared with U-Net, yet the value remains too high for fine-grained tasks. This outcome underscores the challenge posed by complex edges. Row 4 shows condensation in the Petri dish that creates mist-like noise, raising HD95 for model prediction and revealing their shared weakness under noise.

**Robustness evaluation.** Based on the benchmark results, we select SegFormer as the best performer for evaluating model robustness against species-related variations, different types of mycelium culture media, and
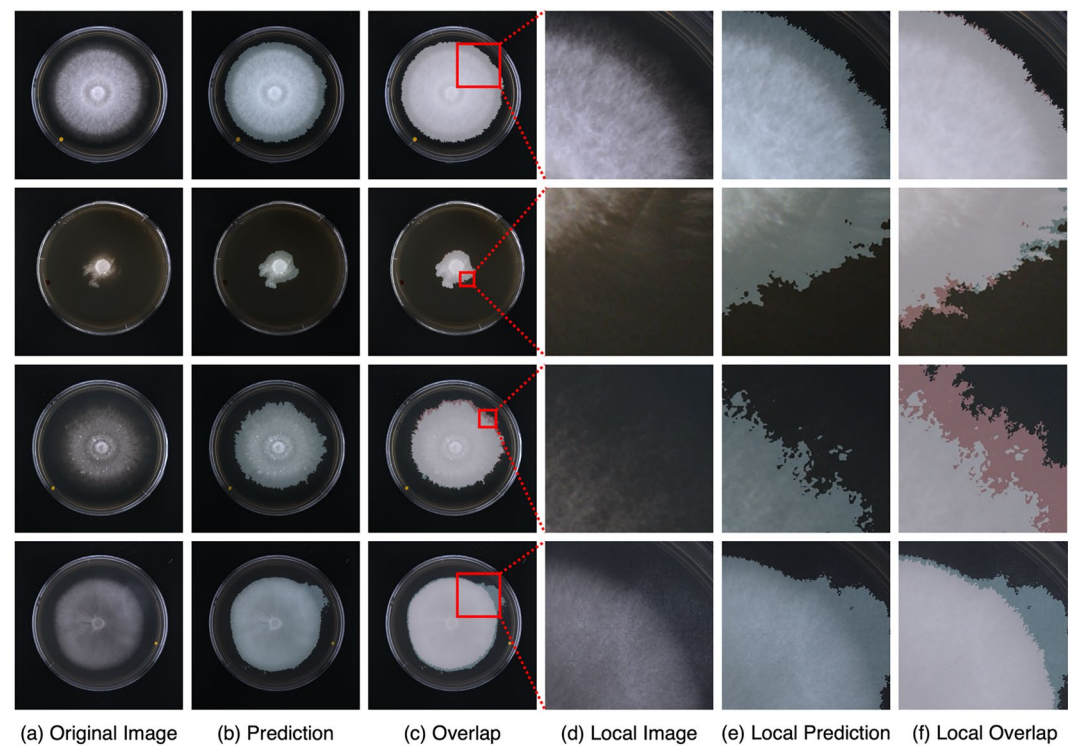
(a) Original Image  (b) Prediction  (c) Overlap  (d) Local Image  (e) Local Prediction  (f) Local Overlap

**Fig. 5** Visualization of baseline model predictions. Blue, red, and white indicate the predicted mask, the ground truth mask, and their overlap, respectively. (**a**) Original image. (**b**) Predicted mask overlaid on the original image. (**c**) Ground truth and prediction overlap overlaid on the original image. (**d**) Magnified crop of the original image. (**e**) Predicted mask overlaid on the magnified crop. (**f**) Ground truth and prediction overlap overlaid on the magnified crop. All panels except (**a**) and (**d**) are produced by blending the corresponding masks with the underlying image using partial transparency.

| Mycelium species | Mycelium F1-score ↑ (%) | Mycelium IoU ↑ (%) | Boundary IoU ↑ (%) | HD95 ↓ (px) | ASSD ↓ (px) |
|---|---|---|---|---|---|
| *Ganoderma sinense* | 96.37 | 92.99 | 18.72 | 97.36 | 21.04 |
| *Pleurotus ostreatus* | 97.67 | 95.44 | 24.50 | 512.53 | 118.29 |
| *Trametes* spp. | 95.93 | 92.18 | 11.50 | 175.08 | 43.77 |

**Table 7.** The results of cross species robustness.

varying culture temperatures. The model maintains consistently stable segmentation performance across species on *Ganoderma sinense*, *Pleurotus ostreatus*, and *Trametes* spp. As shown in Table 7, classic metrics (F1-score and IoU) exceed 92% for all three species. The results of Boundary IoU (from 24.50% to 11.50%) and ASSD (from 21.04 to 118.29) reveal the challenge of boundary-aware segmentation in various fungal species. Figure 6 illustrates model robustness under different types of mycelium culture medium, and varying culture temperatures. In Fig. 6(a), the model maintains robust performance with Mycelium F1-scores and Mycelium IoU both above 93% under 25 °C and 15 °C temperature settings. For boundary-aware metrics, the model performance show highly consistency, with Boundary IoU surrounding 29%, and ASSD value about 16 pixels across two temperature settings. In Fig. 6(b), Mycelium F1-score and Mycelium IoU remain stably above 92% under MYG and PDA culture media settings. The range, in Boundary IoU from 30.50% to 11.50% and in ASSD from 14.10 to 43.77, indicates that boundary aware segmentation under different culture conditions still has substantial room for improvement.

The accuracy of boundary segmentation is crucial for mycelium segmentation research, as it directly affects the quality of studies on core scientific issues, such as quantifying growth patterns, monitoring environmental adaptation, and evaluating physiological responses to different stimuli. In this field, small errors in segmentation boundaries can cause significant inaccuracies in subsequent quantitative analysis results. This highlights the fact that mycelium boundary segmentation poses a highly challenging task.

## Usage Notes
The public release comprises two components: a dataset hosted on Zenodo[24] and a code repository available on GitHub. As for the dataset, researchers could unzip the downloaded archives to obtain two parts data, labeled data and unlabeled data. The labeled data part consists of 'labeled-GL.zip', 'labeled-GS_PO_TS.zip' and 'labeled-MYG_PDA_TEMP.zip'. 'labeled-GL.zip' contains 'trainset' and 'testset' subfolders, which can be used to reproduce the benchmark results or to train, infer, and test custom models. 'labeled-GS_PO_TS.zip' contains
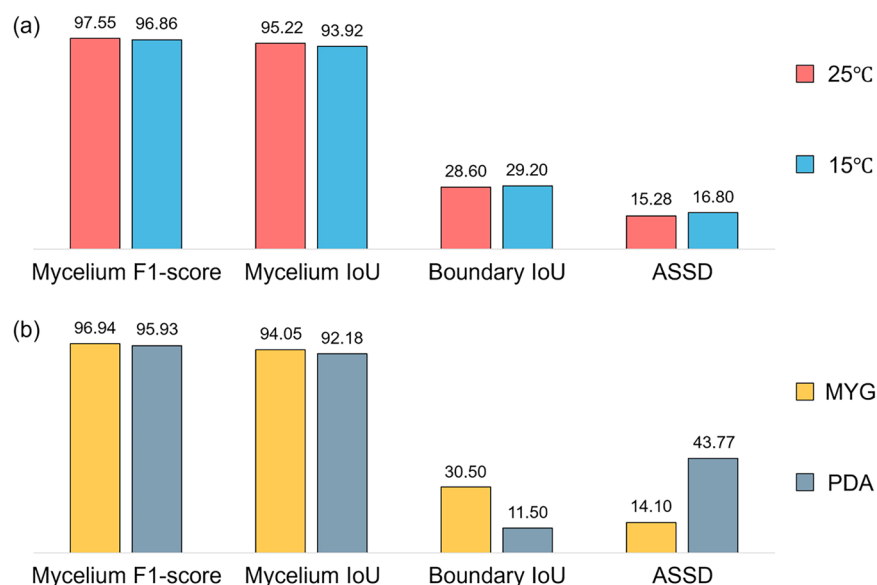
**Fig. 6** Culture condition robustness evaluation results. (**a**) Culture temperature. (**b**) Culture medium.

'GS', 'PO' and 'TS' subfolders, which contains *Ganoderma sinense*, *Pleurotus ostreatus* and *Trametes* spp. labeled samples, separately. These can be used to cross-species robustness test. 'labeled-MYG_PDA_TEMP.zip' contains 'MYG', 'PDA' and 'TEMP15' subfolders, which contains mycelium samples cultured under three conditions: on MYG agar plates, on PDA agar plates, and at 15 °C, respectively. These can be used to environment robustness test. The unlabeled data part provides 19,609 additional images without annotations by eight subfolders. Seven of these named 'unlabeled-GL1' through 'unlabeled-GL7' provides 17,920 *Ganoderma lucidum* images. The remaining named 'unlabeled-GS_PO_TS' provides 1689 images of *Ganoderma sinense*, *Pleurotus ostreatus* and *Trametes* spp. sample. These enable the evaluation of semi-supervised or self-supervised methods for boundary segmentation and supporting various segmentation tasks for further mycelium research. As for the code repository, the repository includes: (a) the 'requirements.txt' file listing all Python dependencies in the form of 'package = =version'; (b) the 'local_configs' folder with default MMSegmentation model, dataset and schedule configurations; (c) the 'mmseg' folder that extends default MMSegmentation with customized evaluation metric function, 'mmseg/core/evaluation/extra_metrics.py', used in this study; (d) the 'mycelium_model' folder containing the dataset configuration file in the path 'mycelium_model/dataset/EPA_mycelium.py', and the model configuration files containing hyper-parameter and module settings for mainstream deep-learning models stored under 'mycelium_model/model'; (e) two shell scripts, 'script_train.sh' and 'script_inference.sh', for training and inference, respectively. Before running the code, the 'data_root' variable in the dataset configuration and the paths in both train and inference scripts should be updated to match their local environment.

## Data availability
The MyceliumSeg dataset used in this study is publicly accessible at Zenodo (https://doi.org/10.5281/zenodo.15224240)[24].

## Code availability
The codes to reproduce the baseline results presented in the Technical Verified section is available at https://github.com/yuanqianguang/MyceliumSeg-benchmark. More information can be found in the associated README.md file.

## References
1. Britton, S. J., Rogers, L. J., White, J. S. & Maskell, D. L. HYPHAEdelity: a quantitative image analysis tool for assessing peripheral whole colony filamentation. *FEMS Yeast Research* **22**, foac060, https://doi.org/10.1093/femsyr/foac060 (2022).
2. Wurster, S. *et al*. Live Monitoring and Analysis of Fungal Growth, Viability, and Mycelial Morphology Using the IncuCyte NeuroTrack Processing Module. *mBio.* **10**, e00673-19, https://doi.org/10.1128/mBio.00673-19 (2019).
3. De Ligne, L. *et al*. Analysis of spatio-temporal fungal growth dynamics under different environmental conditions. *IMA Fungus* **10**, 7, https://doi.org/10.1186/s43008-019-0009-3 (2019).
4. Vidal-Diez De Ulzurrun, G., Huang, T.-Y., Chang, C.-W., Lin, H.-C. & Hsueh, Y.-P. Fungal feature tracker (FFT): A tool for quantitatively characterizing the morphology and growth of filamentous fungi. *PLoS Comput. Biol.* **15**, e1007428, https://doi.org/10.1371/journal.pcbi.1007428 (2019).
5. Hotz, E. C. *et al*. Effect of agar concentration on structure and physiology of fungal hyphal systems. *Journal of Materials Research and Technology* **24**, 7614–7623, https://doi.org/10.1016/j.jmrt.2023.05.013 (2023).

6. Miao, C. *et al*. Semantic Segmentation of Sorghum Using Hyperspectral Data Identifies Genetic Associations. *Plant Phenomics* **2020**, 4216373, https://doi.org/10.34133/2020/4216373 (2020).

7. Dai, W. *et al*. AISOA-SSformer: An Effective Image Segmentation Method for Rice Leaf Disease Based on the Transformer Architecture. *Plant Phenomics* **6**, 0218, https://doi.org/10.34133/plantphenomics.0218 (2024).

8. Yang, X. *et al*. PanicleNeRF: Low-Cost, High-Precision In-Field Phenotyping of Rice Panicles with Smartphone. *Plant Phenomics* **6**, 0279, https://doi.org/10.34133/plantphenomics.0279 (2024).

9. Kapoor, S. & Narayanan, A. Leakage and the reproducibility crisis in machine-learning-based science. *Patterns* **4**, 100804, https://doi.org/10.1016/j.patter.2023.100804 (2023).

10. Upadhyay, A. K. & Bhandari, A. K. Advances in Deep Learning Models for Resolving Medical Image Segmentation Data Scarcity Problem: A Topical Review. *Arch Computat Methods Eng*. **31**, 1701–1719, https://doi.org/10.1007/s11831-023-10028-9 (2024).

11. Subbaswamy, A., Adams, R. & Saria, S. Evaluating model robustness and stability to dataset shift. In *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, Vol. 130, 2611–2619 (PMLR, 2021).

12. Azad, R. *et al*. Medical image segmentation review: The success of u-net. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **46**, 10076–10095, https://doi.org/10.1109/TPAMI.2024.3435571 (2024).

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

14. Madec, S. *et al*. VegAnn, vegetation annotation of multi-crop RGB images acquired under diverse conditions for segmentation. *Scientific Data* **10**, 302, https://doi.org/10.1038/s41597-023-02098-y (2023).

15. Li, Z. *et al*. FPheno2000: Computer Vision-Based Platform for Collection and Intelligent Analysis of Edible and Medicinal Fungal Mycelial Phenotypes. *Journal of Fungal Research*, 1–15, https://doi.org/10.13341/j.jfr.2024.1722 (2024).

16. Ronneberger, O., Fischer, P. & Brox, T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*, 234–241, https://doi.org/10.1007/978-3-319-24574-4_28 (Springer, 2015).

17. Chen, L.-C., Papandreou, G., Schroff, F. & Adam, H. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587* (2017).

18. Xie, E. *et al*. Segformer: simple and efficient design for semantic segmentation with transformers. *Adv. Neur. Inform. Process. Syst.* **34**, 12077–12090 (2021).

19. Cheng, B., Girshick, R., Dollar, P., Berg, A. C. & Kirillov. Boundary IoU: Improving object-centric image segmentation evaluation. In *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 15334–15342, https://doi.org/10.1109/CVPR46437.2021.01508 (2021).

20. Wazir, S. & Kim, D. Rethinking Decoder Design: Improving Biomarker Segmentation Using Depth-to-Space Restoration and Residual Linear Attention. In *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 30861–30871 (2025).

21. Huttenlocher, D. P., Klanderman, G. A. & Rucklidge, W. J. Comparing images using the Hausdorff distance. *IEEE Trans. Pattern Anal. Mach. Intell.* **15**, 850–863, https://doi.org/10.1109/34.232073 (1993).

22. Kavur, A. E. *et al*. CHAOS Challenge - combined (CT-MR) healthy abdominal organ segmentation. *Medical Image Analysis* **69**, 101950, https://doi.org/10.1016/j.media.2020.101950 (2021).

23. Yang, J., Rahardja, S. & Fränti, P. Outlier detection: how to threshold outlier scores? in *Proceedings of the International Conference on Artificial Intelligence, Information Processing and Cloud Computing* 1–6, https://doi.org/10.1145/3371425.3371427 (ACM, Sanya China, 2019).

24. Yuan, Q. *et al*. A Mycelium Dataset with Edge-Precise Annotation for Semantic Segmentation, *Zenodo*, https://doi.org/10.5281/zenodo.15224240 (2025).

25. Huynh, T., Phung, T. V., Stephenson, S. L. & Tran, H. Biological activities and chemical compositions of slime tracks and crude exopolysaccharides isolated from plasmodia of *Physarum polycephalum* and *Physarella oblonga*. *BMC Biotechnol* **17**, 1–10, https://doi.org/10.1186/s12896-017-0398-6 (2017).

26. Steinberg, G., Peñalva, M. A., Riquelme, M., Wösten, H. A. & Harris, S. D. Cell Biology of Hyphal Growth. *Microbiol. Spectr*. **5**, https://doi.org/10.1128/microbiolspec.FUNK-0034-2016 (2017).

27. Dikec, J. *et al*. Hyphal network whole field imaging allows for accurate estimation of anastomosis rates and branching dynamics of the filamentous fungus *Podospora anserina*. *Sci. Rep.* **10**, 3131, https://doi.org/10.1038/s41598-020-57808-y (2020).

28. Loshchilov, I. & Hutter, F. Decoupled weight decay regularization. In *The Seventh International Conference on Learning Representations* https://openreview.net/forum?id=Bkg6RiCqY7 (OpenReview.net, 2019).

29. MMSegmentation Contributors. MMSegmentation: Openmmlab Semantic Segmentation Toolbox and Benchmark. https://github.com/open-mmlab/mmsegmentation (2020).

30. Paszke, A. *et al*. PyTorch: an imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems* **32** (2019).

## Acknowledgements

## Author contributions

Qianguang Yuan: Conceptualization, Methodology, Validation, Investigation, Writing Original Draft, Writing - Review & Editing, Data Annotation, Baseline Model, Visualization, Data Curation. Weizhen Liu: Investigation, Conceptualization, Methodology, Writing - Review & Editing, Resources, Annotation Review, Funding acquisition, Project Administration, Supervision. Yunfei Liu: Data Acquisition, Conceptualization, Validation, Writing Original Draft, Data Annotation, Baseline Model. Pin Li: Data Acquisition, Data Curation. Yuxuan Liu: Data Acquisition, Data Curation. Xiaohui Yuan: Conceptualization, Quality Control, Annotation Review, Writing - Review & Editing. Nanqing Dong: Supervision, Writing - Review & Editing. Shengwu Xiong: Conceptualization, Supervision, Writing - Review & Editing. Yongping Fu: Conceptualization, Supervision, Resources, Funding acquisition, Writing - Review & Editing.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to W.L., S.X. or Y.F.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.