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.

Automated 3D Reconstruction of Mycorrhizal Fungal Structures from Bright-Field Microscopy Using Oriented Bounding Boxes and Focus Modeling

Dalia Y. Domínguez-Jiménez

dalia.dominguez@iit.it

Italian Institute of Technology

Marilena Ronzan

Italian Institute of Technology

Emanuela Del Dottore

University of Southern Denmark

Barbara Mazzolai

Italian Institute of Technology

Research Article

Keywords: Fungal networks, Bioimage Analysis, Microscopy Image Analysis, Network Reconstruction, 3D Reconstruction, Oriented Object Detection

Posted Date: November 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11394217/v1

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

Additional Declarations: No competing interests reported.

Automated 3D Reconstruction of Mycorrhizal Fungal Structures from Bright-Field Microscopy Using Oriented Bounding Boxes and Focus Modeling

Dalia Y. Domínguez-Jiménez [1,2,*], Marilena Ronzan[1], Emanuela Del Dottore [3] and Barbara Mazzolai[1,*]

[1]Bioinspired Soft Robotics Laboratory, Istituto Italiano di Tecnologia, Via Morego 30, Genova 16163, Italy. [2]BioRobotics Institute, Scuola Superiore Sant'Anna, Pisa, Italy. [3]Digital and High-Frequency Electronics, Institute of Mechanical and Electrical Engineering, University of Southern Denmark, Odense, Denmark

* Correspondence: dalia.dominguez@iit.it (D. Domínguez), barbara.mazzolai@iit.it (B. Mazzolai).

Abstract

Background: Fungal networks play a fundamental role in soil ecology and plant symbiosis, yet their three-dimensional characterization remains challenging due to their complex filamentous morphology and the limitations of conventional microscopy image analysis. This work presents a data-driven framework for three-dimensional reconstruction of fungal structures from bright-field microscopy, combining YOLO-based oriented bounding box (OBB) detection with continuous focus modeling for depth estimation. Addressing a problem explored by relatively few existing microscopy-based reconstruction approaches, the framework was developed using images of the arbuscular mycorrhizal fungus *Rhizophagus irregularis*, with additional qualitative applications to other biological structures exhibiting similar filamentous morphologies. **Results:** The OBB detector achieved an mAP50 of 0.60 ± 0.03, with instance-level and spatial-coverage F1-scores of 0.67 ± 0.02 and 0.78 ± 0.02, respectively. To account for the continuous nature of filamentous structures, a composite score combining both evaluation criteria was proposed, yielding S = 0.72 ± 0.01. Focus was

modeled using complementary directional, spectral, and intensity descriptors integrated through PCA, achieving an F1-score of 0.90 on the ×10 validation dataset. The complete framework was applied to a 25-hour time-lapse experiment, enabling repeated three-dimensional reconstruction of the evolving fungal structure. Quantitative spatial evaluation yielded a global pixel-wise recall of 0.93. Furthermore, 75.75% of the focused OBBs overlapped the manual segmentation by at least 20% of their area within an axial tolerance of ±35 μm, as measured by the axial-tolerant overlap rate (ATOR), substantially exceeding spatially randomized controls. **Conclusions:** These results support the feasibility of combining orientation-aware detection and focus modeling for automated three-dimensional characterization of fungal structures, providing a foundation for future investigations of fungal network architecture and growth dynamics.

Keywords

Fungal networks, Bioimage Analysis, Microscopy Image Analysis, Network Reconstruction, 3D Reconstruction, Oriented Object Detection

1. Introduction

Filamentous fungi are ubiquitous in nearly every ecosystem (1,2). Their success in colonizing terrestrial ecosystems is largely attributed to their ability to develop hyphae (3), which grow by apical elongation, branch and fuse through anastomosis to generate a highly interconnected network known as mycelium. This network enables efficient foraging and long-distance transport of nutrients and information signals (4,5). Mycelial networks span a remarkable range of scales, from micrometers to several kilometers, and can persist for

extended periods due to their plastic architecture (6,7), which continuously adapts to spatially and temporally varying environmental conditions and biotic interactions (8).

Within the broad diversity of filamentous fungi, mycorrhizal fungi (MF) represent a key functional group capable of developing mycorrhizal networks that extend outwards from colonized plant roots, through soil, and colonize the roots of multiple neighboring plants, often from different species (9,10). Arbuscular mycorrhizal (AM) fungi, the most widespread group of MF (11), rely entirely on host plants to meet their carbon (C) demand (12). In exchange, their extraradical mycelium explores patchy resource environments, and supplies plants with essential nutrients such as phosphorus (P) and nitrogen (N; 13).

Plants colonized by the same mycorrhizal fungus become interconnected through a Common Mycorrhizal Network (CMN) (5,14), allowing the transmission of nutrients, water, and signaling molecules among members of the network (15,16). Plant responses to AM colonization are highly context-dependent, and CMNs can mediate both competitive and cooperative interactions depending on environmental conditions and species involved (10). Documented benefits to plants include enhanced tolerance to abiotic stresses such as drought (17,18) and protection against biotic pressures, including herbivory, pests, and pathogens (19,20). Although CMNs can strongly influence plant community structure and ecosystem functioning, the mechanisms governing resource exchange, partner choice, and symbiotic outcomes remain unresolved in most cases (10).

These gaps have prompted researchers to study fungal network structure dynamically. Recent work has begun to quantify mycelial topology and morphology over time to understand how fungi shape their architecture and regulate their metabolic demands (21). Morphological and topological traits have also been used to predict functional properties such as transport efficiency, construction cost, and robustness to grazing, which in turn help infer the ecological strategies of different fungal strains (22). Experimental systems still rely heavily on microcosms and controlled environments to circumvent soil opacity, as well as on

image-based methods to reconstruct mycelial morphology (21–24). However, these systems either constrain fungal growth to quasi-two-dimensional environments or rely on two-dimensional image representations, limiting accurate characterization of the intrinsic three-dimensional organization and connectivity of fungal networks.

Several studies have explored X-ray microtomography to investigate the three-dimensional morphology of fungal structures (25,26). These approaches provide direct volumetric information and enable detailed characterization of fungal architecture. However, they are typically applied to dense fungal pellets, require specialized instrumentation and sample preparation, and remain difficult to integrate into high-throughput longitudinal studies of network development. Moreover, while tomography-based methods operate on data that are already volumetric, they remain time consuming and impractical for analyzing the evolution of the network over time. Consequently, microscopy remains the most accessible and widely adopted imaging modality for studying fungal growth and network dynamics.

Through time-lapse microscopy, researchers can monitor fungal development over extended periods while maintaining the temporal resolution necessary to investigate growth, branching, and interactions with the surrounding environment. However, converting these large microscopy datasets into reliable three-dimensional representations of fungal networks remains a major challenge. Existing microscopy-based approaches predominantly rely on image-processing pipelines based on segmentation, skeletonization, and graph extraction (27–29), which often require extensive preprocessing and manual parameter tuning. As a result, automated and scalable reconstruction of fungal networks from microscopy image stacks remains an open problem, motivating the development of automatic computational approaches.

Motivated by these challenges, this work presents an automated, data-driven framework for reconstructing three-dimensional fungal networks from bright-field microscopy image

stacks. In contrast to existing approaches that reconstruct fungal networks from segmented filament masks, the proposed framework represents fungal structures as assemblies of localized, orientation-aware filament fragments detected directly from microscopy images. By combining filament detection, focus-informed depth characterization, and volumetric reconstruction, the framework transforms microscopy observations into three-dimensional representations of mycelial architecture. Specifically, the main contributions are: (i) a newly annotated dataset in which hyphal fragments are represented using oriented bounding boxes (OBBs), providing a compact and orientation-aware description of filamentous structures; (ii) a YOLO-based detector for identifying filament fragments across varying imaging conditions; and (iii) a focus-guided reconstruction strategy that integrates detections across focal planes to recover the three-dimensional organization of fungal networks. Together, these components establish a unified framework for automated fungal network reconstruction and provide a foundation for future analyses of network topology, growth dynamics, and other filamentous systems in biological and medical imaging.

The remainder of this paper is organized as follows. Section 2 reviews existing approaches for the analysis and reconstruction of fungal and filamentous networks, identifying key methodological limitations and motivating the proposed framework. Section 3 presents the proposed methodology, including the experimental setup, dataset construction, orientation-aware detection strategy, focus modeling, and three-dimensional reconstruction pipeline. Section 4 reports the experimental results and validation of the proposed methods. Section 5 discusses the implications, limitations, and potential applications of the framework. Finally, Section 6 summarizes the main conclusions and outlines future research directions.

2. Related works

Filamentous biological networks have been studied across a wide range of disciplines, including fungal ecology, plant–fungus interactions, neuroscience, and vascular imaging.

Depending on the biological question, researchers employ different imaging modalities and computational workflows to extract structural, topological, and dynamic information from these networks. Table 1 summarizes representative image-based approaches across these domains, highlighting their imaging modalities, network representations, extracted descriptors, dimensionality, and degree of automation. This comparison provides context for positioning the proposed framework and identifying the methodological gaps addressed in this work.

For clarity, the reviewed studies are grouped into three categories: biological studies that characterize fungal networks from microscopy images (*Subsection 2.1*), computational methods for fungal network reconstruction and analysis (*Subsection 2.2*), and image-based approaches developed for other filamentous systems (*Subsection 2.3*).

Table 1. Representative image-based approaches for fungal and filamentous network analysis. Methods are compared according to imaging modality, network representation (e.g., segmentation masks, skeletons, graphs, or filament fragments), temporal scope (static, dynamic, or multi-timepoint analysis), dimensionality of the reconstructed network (2D or 3D), and degree of data-driven processing. Temporal scope is classified as static when a single time point is analyzed, multi-timepoint when multiple time points are analyzed or reconstructed independently, and dynamic when explicit temporal correspondence is established between successive time points to track structures or events over time. Data-driven processing is classified as No for predominantly handcrafted methods, Partial for hybrid approaches, and Yes for predominantly data-driven methods

Domain	Imaging modality	Representation	Temporal Scope	2D/3D	Data-Driven	Ref.

Cytoskeleton	Super-resolution microscopy	Skeleton + Filament Graph	Static	2D	No	(30)
Microorganisms	Bright-field microscopy	Segmentation Mask + Skeleton	Dynamic	2D	No	(31)
Slime molds	Bright-field microscopy	Segmentation Mask	Dynamic	2D	No	(29)
Neurons	Volumetric microscopy	Tree Graph	Static	3D	No	(32)
	Confocal microscopy	Instance Segmentation Mask	Static	3D	Partial	(33)
Leaf venation	Light microscopy	Segmentation Mask + Graph	Static	2D	Yes	(32)
Retinal vessels	Fundus photography	Segmentation Mask	Static	2D	Yes	(34,35)
Fungal pellets	µCT	Segmentation + Skeleton	Static	3D	No	(25)
			Dynamic	3D	Partial	(26)
Fungal networks	Bright-field microscopy	Segmentation Mask + Skeleton + Graph	Static	2D	No	(27,36)
		Ridge Map + Skeleton + Graph	Static	2D	No	(37)
		Segmentation Mask + Graph	Dynamic	2D	No	(28)
	Bright-field microscopy (Z-stack)	Segmentation Mask + Skeleton + Graph	Static	2D	No	(38)
		Skeleton + Graph	Multi-timepoint	3D	No	(37)
		OBB-based filament representation	Multi-timepoint	3D	Yes	Ours

2.1 Quantitative Characterization of Fungal Network Architecture

Imaging has become an essential tool for studying fungal networks because their morphology and function are inherently spatial and dynamic. Across experimental systems, image analysis typically serves as the quantitative measurement layer that transforms complex mycelial architectures into descriptors suitable for biological interpretation rather than constituting the primary scientific contribution itself. Spanning both microscopic and

150 macroscopic scales, from hyphal-level observations to cord-forming networks, these
151 measurements enable quantitative characterization of mycelial expansion, biomass
152 distribution, and network organization under different environmental conditions and
153 biological interactions (21,22,39,40).

154 The resulting image-derived measurements support a broad range of biological
155 questions. They are used to characterize colony development and early network formation
156 (23,24), growth responses to environmental conditions such as temperature and humidity
157 (39), plant–fungus colonization dynamics and resource exchange (21), and the relationship
158 between network architecture, transport, and ecological strategies (7,22). Together, these
159 studies demonstrate that quantitative image analysis underpins our understanding of how
160 fungal networks develop, adapt, and function across different biological contexts.

161 Accordingly, the descriptors extracted from fungal images can be broadly grouped into
162 three complementary categories. Morphological descriptors quantify colony expansion and
163 local hyphal organization through measurements such as colony area, total hyphal length,
164 colony diameter, tip number, hyphal segment length, branching characteristics, and hyphal
165 growth unit (23,24,39). Topological descriptors represent the mycelium as a graph to
166 characterize node and edge organization, degree, connectivity, loops, cross-connectivity, and
167 other network properties associated with architectural organization (7,22,23). Dynamic
168 descriptors capture the temporal evolution of fungal networks through growth and branching
169 dynamics, anastomosis events, colonization fronts, transport efficiency, construction cost,
170 robustness, and responses to environmental perturbations (21,22,24,39).

171 Despite their biological value, the reliability of these measurements depends
172 fundamentally on the underlying image-analysis methodology. Existing workflows remain
173 largely tailored to specific experimental settings and predominantly rely on two-dimensional
174 representations, which can obscure the distinction between filament crossings and true
175 anastomoses while biasing connectivity-related measurements (22,23). Their performance is

further affected by variations in image quality, heterogeneous backgrounds, and dense filament organization, frequently requiring manual correction, limiting reproducibility and scalability, and hindering accurate characterization of the intrinsic three-dimensional organization and connectivity of complex fungal networks (5,7,39). These challenges motivate the development of dedicated computational methods for fungal network reconstruction and analysis, which are discussed in the following sections.

2.2 Computational Methods for Fungal Network Reconstruction

Early and widely adopted image-processing approaches for fungal network analysis follow a segmentation–skeletonization paradigm, in which filamentous structures are extracted as binary masks, reduced to centerlines, and subsequently converted into graph representations (28,36,38). These graphs enable quantification of morphological and topological descriptors, including hyphal length, branching statistics, network connectivity, tortuosity, and transport-related measures, and have been successfully applied to fungal networks grown in soil microcosms and planar agar cultures (27,24,35). However, their performance remains highly dependent on image preprocessing, segmentation quality, and many user-defined parameters. Consequently, robustness decreases under heterogeneous, noisy, or low-contrast imaging conditions, requiring careful calibration of filtering, thresholding, and pruning operations that limit scalability across experimental settings (27,28,36).

In parallel with whole-network reconstruction pipelines, several approaches have focused on time-resolved analysis of fungal growth, prioritizing dynamic descriptors over full topological reconstruction (29,35–37). These methods typically operate on time-lapse image sequences and track either individual hyphal tips, regions of interest, or simplified

198 backbone representations to estimate elongation rates, branching events, and growth
199 persistence. Representative examples include dynamic segmentation and backbone tracking
200 frameworks (31), ROI-based and area-driven analyses of spore germination and early hyphal
201 development (42), and continuous video microscopy systems designed to follow single
202 hyphae in real time (43). Although these approaches provide valuable insights into growth
203 dynamics, they generally restrict analysis to sparse or isolated filaments and exclude
204 overlapping or interacting structures, limiting their applicability to whole-network analysis.
205 Beyond segmentation-based pipelines, ridge-based representations have been proposed
206 to better preserve filament continuity and local geometry. In slime mold transport networks,
207 ridge enhancement combined with skeletonization and graph analysis has been used to
208 characterize network organization, dynamics, and transport properties while maintaining
209 biologically meaningful connectivity (29). Similar strategies have subsequently been applied
210 to fungal networks, revealing a key limitation of two-dimensional imaging: branching events
211 and filament crossings often produce similar image patterns, making the underlying network
212 connectivity difficult to infer from a single focal plane (44). To address this limitation,
213 volumetric approaches have reconstructed three-dimensional filament graphs from z-stack
214 images by linking skeleton segments across focal planes using spatial proximity and
215 orientation consistency (37). Together, these studies highlight the importance of preserving
216 filament geometry and motivate reconstruction methods capable of recovering the intrinsic
217 three-dimensional organization of fungal networks.
218 Explicit three-dimensional reconstruction of fungal networks from microscopy data has
219 only recently begun to emerge. To the best of our knowledge, Sten et al. (37) presented the
220 only microscopy-based framework capable of reconstructing fungal architectures from z-
221 stack images while tracking their evolution across multiple time points. Their approach
222 combines Shape-from-Focus (SFF) depth estimation with skeleton-based graph

reconstruction to recover volumetric network structure. However, SFF accuracy depends on the selected focus measure operator, local image texture, evaluation-window size, and image noise (45–47), while the final reconstruction remains sensitive to skeleton extraction and inter-slice association. Motivated by these limitations, the proposed framework adopts a learned filament representation based on oriented bounding boxes (OBBs), which explicitly encode the position, orientation, and spatial extent of local filament fragments. By detecting filament-centered regions directly from image data and integrating complementary focus information at the fragment level across depth, the method reduces dependence on predefined evaluation windows and skeleton-based reconstruction, while providing a scalable framework for three-dimensional analysis of fungal and other filamentous biological networks.

2.3 Data-Driven Approaches for Filamentous Structures

Recent advances in deep learning (DL) have shifted filamentous image analysis from rule-based pipelines based on handcrafted filtering, thresholding, and skeletonization toward data-driven models that learn image features directly from annotated data. As summarized in a systematic review of filament tracing and tracking (48), most DL approaches formulate the problem as semantic or instance segmentation, producing dense pixel-wise masks that subsequently support tracing, graph extraction, or quantitative analysis. While these methods substantially reduce manual parameter tuning during inference, they generally require extensive pixel-level annotations for training.

Segmentation-based DL has been successfully applied across a broad range of filamentous biological structures, including retinal vessels, volumetric vascular networks, cytoskeletal filaments, and plant root systems (44–48). Despite differences in application, these methods share a common workflow in which segmentation serves as the initial representation, followed by additional processing to recover connectivity, topology, or

248 individual filament instances. In retinal imaging, connectivity-aware extensions combine
249 learned image features with graph-based reasoning to preserve long-range vessel structure
250 under low-contrast conditions (49), whereas in two-photon microscopy, generalizable models
251 emphasize robustness to depth-dependent signal degradation and heterogeneous acquisition
252 settings (50). Similar challenges arise in cytoskeletal filament analysis, where semantic
253 segmentation alone is insufficient for downstream quantification, motivating instance-
254 oriented and orientation-aware formulations to separate individual filaments at intersections
255 (51). In plant phenotyping, segmentation models have been adopted for root and related
256 structures (52), alongside interactive or corrective annotation strategies designed to reduce
257 labeling effort while maintaining robustness across variable backgrounds and imaging
258 conditions (53).

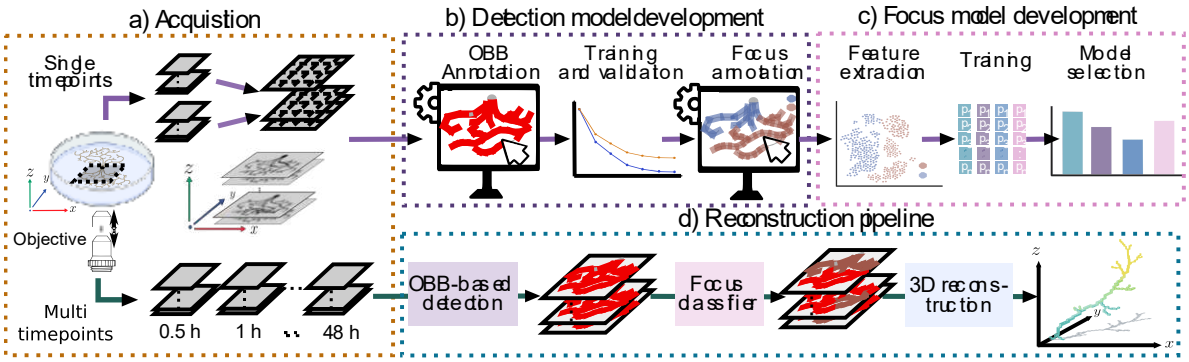
259 Despite these advances, segmentation-based methods remain constrained by the need for
260 dense annotations and by their implicit representation of filament geometry, where
261 orientation and continuity must be inferred through subsequent processing. Furthermore, DL
262 methods have not yet been reported for the segmentation, tracing, or tracking of mycelial
263 networks. These limitations motivate the exploration of alternative learning formulations that
264 explicitly preserve the local geometric information required for downstream reconstruction.
265 Among these, object detection represents image content through sparse geometric primitives
266 rather than dense pixel masks. When formulated using OBBs, it explicitly encodes the
267 position, orientation, and spatial extent of local filament fragments. These descriptors directly
268 support fragment association, graph extraction, and correspondence across adjacent image
269 slices or time points, while orientation provides an additional geometric cue for resolving
270 ambiguities arising from filament crossings, branching events, and anastomoses.
271 Consequently, representing filamentous structures as assemblies of oriented fragments offers
272 a compact geometric representation naturally suited to three-dimensional network
273 reconstruction.

Object detection has become a well-established computer vision paradigm for localizing and classifying objects in images (54). While conventional detectors rely on axis-aligned bounding boxes, OBB formulations extend this representation by explicitly encoding object orientation, making them particularly suitable for elongated and anisotropic structures. Originally developed for applications such as aerial imagery and scene text analysis (55,56), OBB detection has since been incorporated into modern end-to-end architectures, including recent YOLO-based detectors, enabling accurate and scalable orientation-aware detection (57). However, to the best of our knowledge, OBB-based detection has not previously been explored for biological filamentous structures, particularly fungal filament networks, nor has it been explored as a geometric representation for three-dimensional reconstruction from bright-field microscopy.

3. Methodology

This section presents the proposed data-driven framework for reconstructing three-dimensional fungal networks from time-lapse bright-field microscopy. As illustrated in Figure 1, the methodology comprises two complementary phases. The first (Fig. 1b,c) focuses on framework development, including dataset construction, annotation, and the training and validation of the OBB detection and focus-classification models using independently acquired single-time-point Z-stacks. The second (Fig. 1d) corresponds to framework deployment, where the trained models are applied to time-lapse microscopy data and their outputs are integrated to reconstruct the three-dimensional morphology of fungal networks. The following subsections describe each stage of the proposed framework in detail.

Figure 1. Overview of the proposed data-driven framework for reconstructing filamentous fungal networks: (a). The methodology comprises two complementary phases: framework development (b–c), including dataset construction, model development, and validation, and framework deployment (d), where the trained models are integrated for three-dimensional reconstruction.



3.1 Biological experiment and sample preparation

Arbuscular mycorrhizal fungus *Rhizophagus irregularis* (MUCL 41833) was provided by the Glomeromycota IN vitro COllection (GINCO) at the Université Catholique de Louvain, Belgium, and cultivated at the center of Petri dishes containing MSR medium. Cultures were maintained at 27 °C in the dark to allow mycelial development. Cultures were maintained for approximately three weeks prior to imaging. Multiple independent cultures were prepared following the same protocol. For imaging experiments, Petri dishes were removed directly from the incubation chamber and placed on the microscope stage without further manipulation or sample preparation.

3.2 Image acquisition and dataset overview

Image acquisition was performed using a Nikon Eclipse Ti2-E inverted microscope operated in bright-field transmission mode. Images were acquired using three objective magnifications (×4, ×10, and ×20). All images were recorded as single-channel 12-bit grayscale images with a spatial resolution of 2928 × 2928 pixels. For each sample, image data were collected as Z-stacks spanning the vertical extent of the fungal network. Acquisition began at the first focal plane where hyphae appeared in focus and continued along the optical axis until the structures became completely defocused. Consecutive focal planes were acquired with a constant Z-step of 5 μm, producing image stacks that captured the full depth of the mycelial network.

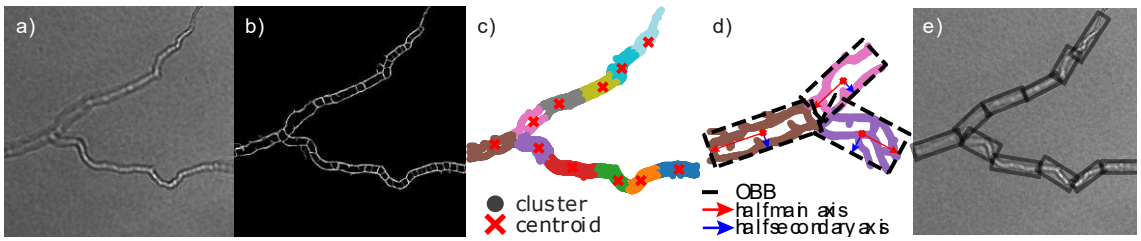
Datasets acquired at the three objective magnifications served distinct roles in the development and evaluation of the proposed framework. Single-time-point Z-stacks acquired with the ×4 and ×20 objectives were used to construct the annotated datasets for both OBB detection and focus classification. These objectives represent the lower and upper bounds of the imaging magnifications employed in this study and were selected to expose model development to substantial variations in filament scale and appearance. In contrast, the intermediate ×10 objective was used for a separate time-lapse experiment, providing data for focus-model validation and calibration as well as multi-timepoint 3D reconstruction. The microscopy datasets and associated annotations used in this study are publicly available at Zenodo (58).

The annotated dataset comprised 21 single-time-point Z-stacks, each acquired from an independent fungal culture. These stacks were selected to represent typical imaging conditions encountered *in vitro* fungal growth experiments, including gradual focus variation across depth, heterogeneous hyphal density, and frequent filament crossings and overlaps. A separate multi-timepoint experiment was performed on a separate fungal culture. The first time point was used for focus-model validation and threshold calibration, while the complete sequence was used for multi-timepoint three-dimensional reconstruction.

3.3 OBB dataset creation

This section describes the procedure used to generate OBB annotations from raw microscopy images. The pipeline combines a validated ridge-based preprocessing method with clustering, geometric estimation, and manual refinement, and was designed to produce consistent, orientation-aware annotations for elongated filamentous structures (Fig. 2).

Figure 2. OBB dataset creation pipeline: a) Original microscopy image slice. b) Composite ridge-based edge map obtained by multi-scale filtering, highlighting elongated filamentous structures. c) Cluster assignment of sampled ridge centroids, where colored points indicate local filament fragments and red crosses denote cluster centroids. d) OBB estimation for selected clusters using PCA; dashed outlines indicate OBBs, while red and blue arrows represent the half-length (major) and half-width (minor) axes, respectively. e) Final OBB representation overlaid on the original image after manual refinement using a custom annotation interface.



3.3.1 Ridge-based preprocessing

To initialize the annotation process, the filament-detection algorithm previously developed and validated by Sten et al. (37) was applied to each image slice. The original MATLAB implementation, publicly available at https://gitlab.iit.it/oscar.sten/mycelium-network-analysis-tool-3d, was used in this study. The method enhances elongated ridge-like structures using a steerable second-order Gaussian filter and is controlled by two scale-dependent parameters: the spatial standard deviation of the filter kernel (σ) and the minimum

branch length. In the present study, the detection stage was used exclusively to generate binary ridge-derived representations for annotation initialization. The minimum branch length was fixed at 100 pixels, whereas σ was adjusted according to the apparent filament width, using values within the range of 2–10 pixels. When necessary, multiple σ values were applied to the same image, and the resulting binary masks were combined by pixel-wise summation to produce a composite ridge map capturing filamentous structures across multiple scales (Fig. 2b). When required, the resulting binary masks were manually refined using MATLAB's *Image Segmenter* app to correct incomplete ridge responses, remove spurious detections, and improve edge continuity prior to centroid sampling. These preprocessing steps were performed in a semi-automatic and qualitative manner to improve edge visibility while accommodating variations in filament morphology and image appearance.

3.3.2 Centroid sampling and clustering

Following ridge-based preprocessing, a binary segmentation mask was constructed for each Z-stack by superimposing the ridge-derived masks from all image slices. Morphological operations and manual corrections were subsequently applied to improve filament continuity, remove spurious structures, and obtain a coherent representation of the fungal network. Candidate points were then generated through an iterative sampling procedure within the binary segmentation mask. Only points located inside the segmented regions were retained, and a minimum spacing of 50 pixels was enforced between neighboring points to obtain an approximately uniform spatial distribution. The sampling procedure was repeated until an empirical sampling density of approximately one point per 200 foreground pixels was reached.

All sampling results were visually inspected using a custom graphical user interface (GUI) developed in Python. When necessary, individual sampled points were manually

added, removed, or repositioned to improve the representation of highly curved filaments, dense network regions, or areas where the automatic sampling was insufficient. Finally, the corrected set of sampled points was grouped using *k*-means clustering (Fig. 2c), where each cluster represented a localized filament fragment suitable for approximation by a single OBB. The number of clusters was selected according to the resulting distribution of sampled points, allowing the clustering process to adapt naturally to variations in filament geometry and local network density.

3.3.3 OBB computation via PCA

An OBB can be represented either by its center coordinates, dimensions, and orientation, (c_x, c_y, w, h, θ), or equivalently by the coordinates of its four vertices. Since YOLO-OBB requires point-based annotations, the estimated geometric parameters were analytically transformed into the normalized coordinates of the four OBB vertices. Principal component analysis (PCA) was applied to each centroid cluster to estimate the dominant orientation of the filament fragment. The cluster centroid defined the OBB center, $\mathbf{c} = (c_x, c_y)$, while the first principal component defined the orientation angle θ. Projections of the sampled points onto the principal axes yielded the OBB length (w) and thickness (h, minor axis).

The OBB corners were first expressed relative to the box center. They were then rotated by the estimated orientation and translated to the cluster centroid to obtain the final vertex coordinates, yielding the four OBB vertices through the following matrix transformation:

$$\mathbf{P} = \mathbf{1}\mathbf{c}^T + \mathbf{Q}\mathbf{R}^T \quad (1)$$

where ($\mathbf{P} \in \mathbb{R}^{4\times 2}$) contains the coordinates of the four OBB vertices, ($\mathbf{c} = [c_x, c_y]^T$) denotes the box center, and (1) is a (4×1) vector of ones. The rotation matrix $\mathbf{R} = [\cos\theta \quad -\sin\theta \; ; \; \sin\theta \quad \cos\theta]$, defines the OBB orientation, while $\mathbf{Q} = \frac{1}{2}[w \quad h \; ; \; w \quad -h; \; -w \quad -h; \; -w \quad h]$ contains the local coordinates of the four OBB

corners expressed as signed half-length and half-thickness offsets from the box center (Fig. 2d).

Finally, the vertex coordinates were normalized with respect to the image dimensions, $x_i' = \frac{x_i}{W}, y_i' = \frac{y_i}{H}, i = 1, \ldots, 4$. where W and H denote the image width and height, respectively. The resulting normalized vertices constituted the point-based annotations used for training the YOLO-OBB detection model.

3.3.4 Manual correction and annotation interface

Following automatic OBB estimation, a second custom graphical user interface (GUI) was developed in Python to review and refine the generated annotations. The interface enabled interactive visualization of individual image slices with the estimated OBBs overlaid on the original microscopy images. Using mouse and keyboard controls, annotators could modify the position, orientation, and dimensions of each OBB, as well as add, delete, or replace annotations when necessary. This final refinement stage was performed to resolve ambiguities arising from filament crossings, variable image contrast, partial ridge responses, or inaccuracies introduced during the automatic clustering and PCA estimation. The resulting manually validated OBB annotations constituted the final dataset used for training the detection model.

3.4 OBB-based detection model

An oriented object detection model was trained to identify local filament fragments in individual image slices using OBBs. The model is based on the YOLO-OBB framework implemented in the Ultralytics library (59), which extends single-stage YOLO detectors to directly regress OBBs within a unified architecture. In this study, the lightweight YOLO11n-OBB variant was selected to balance detection accuracy and computational efficiency for large microscopy datasets.

Training data consisted of the OBB annotations described in Section 3.3. Each image slice was subdivided into 732 × 732-pixel tiles for training, with empty tiles excluded. During model preprocessing, the tiles were resized by the Ultralytics pipeline to the configured network input resolution of 736 × 736 pixels. The model was trained in a single-class configuration, as all detections correspond to filamentous structures. Models were trained for up to 100 epochs using a batch size of 8. Model parameters were initialized from random weights, and deterministic training was ensured through fixed random seeds. To improve robustness to variations in filament orientation, scale, and partial occlusion, geometric data augmentation was applied during training, including rotations, flips, scaling, translations, Mosaic, and CutMix augmentations. Color-based augmentations were disabled because all images were acquired as single-channel grayscale images.

Model performance was assessed using five-fold cross-validation, where annotated mosaic samples from all annotated Z-stacks were pooled and randomly assigned to the training and validation folds. Conventional object detection metrics were first computed using one-to-one matching between predicted and ground-truth OBBs. However, these metrics alone are often overly restrictive for filamentous structures, where multiple detections may correspond to different portions of the same annotated filament fragment.

To complement the instance-level evaluation, a coverage-based metric was also employed. Each predicted OBB was represented as an anisotropic two-dimensional Gaussian distribution aligned with its centroid, orientation, length, and thickness, following the geometric representation proposed for OBBs (60). The Gaussian responses from all detections were superimposed to generate a continuous filament coverage map. The resulting density field was thresholded to obtain a binary support region and compared with the ground-truth annotations using segmentation-based precision, recall, and F1 score. Unlike the instance-level evaluation, this measure quantifies how accurately the predicted detections

recover the spatial support of the underlying filament network, independently of the number of predicted OBBs.

Overall detection performance was summarized using a balanced composite score S combining the instance-level and coverage-based evaluations,

$$S = \alpha F1_{det} + (1 - \alpha) F1_{cov} \quad (2)$$

where $F1_{\mathrm{det}}$ denotes the instance-level detection F1 score, $F1_{\mathrm{cov}}$ denotes the coverage-based F1 score, and $\alpha = 0.5$. Precision and recall are reported separately for both evaluation schemes.

Detection performance depends on two inference-time parameters: the confidence threshold used to retain predictions and the intersection-over-union (IoU) threshold used for non-maximum suppression (NMS) of overlapping predicted OBBs. To select these parameters objectively, the composite score S was evaluated over a grid of IoU thresholds ranging from 0.50 to 0.95 (step 0.05) and 1000 uniformly sampled confidence thresholds in the interval $([0,1])$.The operating point was defined as the parameter pair maximizing the mean composite score across the validation folds and was subsequently fixed for all reported experiments.

After cross-validation, a final detector was trained using the complete annotated dataset and deployed with the selected operating thresholds. Generalization was evaluated on previously unseen microscopy images acquired under different imaging conditions, including images collected with a ×10 objective lens, as well as on additional filamentous datasets comprising plant root microscopy, retinal vessel images from the DRIVE dataset (61), and cerebellar climbing fiber microscopy stacks (62). All external images were converted to grayscale and processed using the same 736 × 736-pixel tiled inference strategy used for model deployment. Predicted detections from individual tiles were subsequently transformed back into the original image coordinates to reconstruct the complete detection results.

3.5. Focus Modeling and Depth-Aware Filtering

Limited depth of field in bright-field microscopy causes filament structures to appear across multiple adjacent slices within a Z-stack with varying degrees of focus. Although these observations may correspond to the same underlying filament, only a subset provides a sharp representation near its focal plane, whereas neighboring slices contain progressively defocused appearances. Retaining all detections may introduce artificial continuity across slices or spurious branches, complicating the interpretation of filament topology. To mitigate these effects, a focus-aware filtering stage was introduced to identify the OBBs that best represent each filament fragment before three-dimensional reconstruction.

Focused filament fragments exhibit characteristic visual properties, including increased contrast, sharper intensity transitions, and greater high-frequency content, as widely reported in the literature on focus measure operators (52–54). Accordingly, each detected OBB was characterized by a set of 18 complementary focus descriptors (Table 2), computed from the image patch enclosed by the corresponding bounding box. These descriptors comprise four groups: classical focus measures, spectral features, intensity descriptors, and directional derivatives. The directional descriptors exploit the orientation provided by the OBB to quantify anisotropic intensity variations along and across the filament, extending conventional isotropic focus operators. Mathematical definitions of all descriptors are provided in the Additional file 1: Table S1.

Table 2. Focus-related features extracted from OBB-associated image patches. The evaluated descriptors are grouped into four categories: classical focus measures, spectral features, intensity descriptors, and directional derivatives. For each feature, the corresponding operator name and symbol used throughout the manuscript are reported.

Category	Focus operator	Sym.	Category	Focus Operator	Sym.
	Laplacian variance	$\sigma^2_{\nabla^2 I}$		Local variance energy	$E_{var,loc}$

Classical focus measures	Modified Laplacian	ML	Intensity descriptors	Mean intensity	μ_I
	Sum Modified Laplacian	SML		Intensity energy	E_I
Spectral features	Fourier energy	E_F	Directional derivatives	Mean absolute normal second derivative	$\mu_{\lvert\partial^2 I/\partial n\rvert}$
	High frequency energy (15%, 30%)	$E_{HF}^{(\%)}$		Normal second-derivative energy	$E_{\partial^2 I/\partial n^2}$
	High frequency ratio (15%, 30%)	$R_{HF}^{(\%)}$		Mean absolute tangential second derivative	$\mu_{\lvert\partial^2 I/\partial t^2\rvert}$
	High-to-low frequency ratio (15%)	$\rho_{HF}^{(\%)}$		Tangential second-derivative energy	$E_{\partial^2 I/\partial t^2}$
				Mean absolute normal gradient	$\mu_{\lvert\partial I/\partial n\rvert}$
				Tangential gradient energy	$E_{\partial I/\partial t}$

Focus was modeled as a continuous quantity describing the degree of sharpness of each filament fragment across the Z-stack. Accordingly, each OBB was assigned a scalar focus score (s_i), with larger values indicating greater focus. To guide the development and evaluation of this representation, a subset of OBB-associated image patches was manually annotated as either in focus (class 1) or out of focus (class 0), based on visual sharpness and consistency with filament continuity across adjacent slices. The annotated subset was balanced independently for each acquisition objective, with equal numbers of focused and defocused samples. It comprised 5,878 ×4 patches, 7,950 ×20 patches, and 2,882 ×10 patches, corresponding to 35.2%, 47.6%, and 17.2% of the total annotated focus dataset, respectively. The ×4 and ×20 samples constituted the model-development dataset, whereas the ×10 samples were held out from model fitting and used for validation, model selection, and operating-threshold calibration.

The annotated development dataset was used to identify a compact set of complementary descriptors for characterizing focus. Feature relevance was first assessed using standardized

logistic regression with five-fold stratified cross-validation. The discriminative contribution and stability of each descriptor were characterized from the distribution of the absolute regression coefficients across folds. Redundancy among descriptors was subsequently assessed using the absolute pairwise Pearson correlation coefficients computed on the model-development data. Based on feature relevance and pairwise correlations, a pool of seven representative descriptors was retained to preserve complementary focus-related information across feature categories while reducing redundancy. Candidate feature triplets were subsequently constructed from this pool and evaluated using the PCA-based focus-score formulation described below.

Rather than directly training a binary classifier, the selected descriptors were integrated into a continuous latent representation of focus. PCA with two components ($K = 2$) was applied to the standardized descriptors to capture their dominant joint modes of variation. The resulting two-dimensional latent representation was subsequently reduced to a scalar focus score by combining the oriented principal-component projections. This formulation models focus as a gradual property of filament appearance rather than an intrinsically binary state.

For each candidate descriptor triplet, the PCA-based focus score was constructed as follows. Let $\mathbf{x}_\mathrm{i} = [\mathrm{x_{i1}}, \mathrm{x_{i2}}, \mathrm{x_{i3}}]$ denote the feature vector associated with the i-th OBB. Each feature was standardized using statistics estimated from the development set,

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j} \tag{3}$$

where μ_j and σ_j denote the mean and standard deviation of feature j, respectively, estimated from the training data. The standardized feature vectors were projected onto a two-component PCA representation,

$$y_{ik} = \mathbf{z}_i^\top \mathbf{v}_k \tag{4}$$

where $\boldsymbol{v}_k$ denotes the loading vector of the k-th principal component. Because the sign of a principal component is arbitrary, each component projection was oriented using the annotated training labels such that larger values were associated with the focused class,

$$\tilde{y}_{ik} = sign_k y_{ik} \tag{5}$$

where $(sign_k \in -1, +1)$ is determined by comparing the mean component projections of the focused and defocused training samples. The continuous focus score was then defined as a weighted combination of the oriented component projections, with weights ω_k obtained by normalizing the explained-variance ratios of the $K = 2$ PCA components such that $\sum_{k=1}^{K} \omega_k = 1$. Thus, components explaining a greater proportion of the represented feature variance contributed more strongly to the score,

$$s_i = \sum_{k=1}^{K} \omega_k \ \tilde{y}_{ik} \tag{6}$$

The resulting scalar score was standardized using the mean and standard deviation of the raw focus scores in the training data,

$$\hat{s}_i = \frac{s_i - \mu_s}{\sigma_s} \tag{7}$$

where μ_s and σ_s denote the training-score mean and standard deviation, respectively. Finally, each OBB was classified as focused or defocused according to

$$\hat{y}_i = \begin{cases} 1 & \text{if } \hat{s}_i \geq \tau \\ 0 & \text{otherwise,} \end{cases} \tag{8}$$

For each candidate triplet, a decision threshold τ was estimated from the model-development data by maximizing the F1 score. The fitted PCA representation, including the feature-standardization parameters, PCA loadings, component orientations, score weights, and development-set decision threshold, was then applied to the ×10 dataset without refitting. Performance on the ×10 data was used to compare candidate focus representations and select the final triplet. After selection, the continuous focus representation was kept fixed, while its

operating threshold τ was recalibrated for the ×10 acquisition conditions by maximizing the F1 score on the ×10 focus scores. The resulting ×10-specific threshold was subsequently kept fixed for all time points of the ×10 time-lapse experiment and used for depth-aware filtering during three-dimensional reconstruction.

3.6 3D reconstruction and time-lapse integration

Time-lapse imaging was performed on a separate fungal culture using repeated Z-stack acquisitions over a 25-hour period. Image stacks were acquired at approximately 30-min intervals using a ×10 objective, yielding 48 time points over the acquisition period. Each time point consisted of a Z-stack containing 165 optical sections acquired with a constant axial spacing of 5 μm, covering the full depth of the fungal network. The first time point was used as the ×10 validation and calibration dataset described in *Section 3.5* to evaluate the fixed PCA-based focus representation and calibrate the ×10 operating threshold.

For each time point, every image slice of the corresponding Z-stack was processed independently using the trained YOLO11n-OBB detector. For inference, each 2928 × 2928-pixel slice was padded along its bottom and right boundaries to 2944 × 2944 pixels and partitioned into sixteen non-overlapping 736 × 736-pixel tiles. The tiles were processed independently using the confidence and intersection-over-union thresholds selected during model validation (Section 3.4). The resulting detections were transformed back into the original image coordinates and merged to recover the complete set of OBBs for each slice.

The detected OBBs were then processed using the fixed PCA-based focus representation and the calibrated ×10 operating threshold described in *Section 3.5.* Only detections classified as in focus were retained, thereby reducing depth ambiguity by discarding defocused appearances of filament fragments detected across adjacent optical sections. The filtered OBBs were subsequently embedded into three-dimensional space using their centroid coordinates (x, y), the corresponding slice position (z), and their estimated geometric

parameters (width, height, and orientation). Image coordinates were converted into physical units using a lateral calibration factor of $0.60\ \mu m\ pixel^{-1}$, while the axial coordinate was determined from the slice index and the known Z-step of $5\ \mu m$. Each retained OBB therefore represents the spatial location, orientation, and local extent of a filament fragment within the reconstructed volume. Applying this procedure independently to each Z-stack produced a sequence of three-dimensional reconstructions across the 48 acquisition time points, providing a multi-timepoint representation of fungal network development.

3.6.1 Quantitative spatial evaluation

To quantitatively assess the spatial correspondence between the reconstructed filament distribution and manual annotations, two complementary evaluation measures were employed. First, pixel-wise recall was calculated as the fraction of manually annotated pixels covered by the union of focused OBBs at their corresponding optical sections,

$$Recall_{pixel} = \frac{\sum_{z\epsilon\mathbb{Z}}|G_z \cap M_z|}{\sum_{z\epsilon\mathbb{Z}}|G_z|} \quad (9)$$

where G_z denotes the binary manual segmentation mask at optical section z, M_z represents the union of rasterized focused OBBs at the same section, and Z denotes the set of evaluated optical sections. The operator $|\cdot|$ indicates the number of pixels belonging to a binary region. Recall was computed globally by aggregating pixel counts across all optical sections.

The second measure was introduced to account for differences in the focal-plane localization of filament fragments between predictions and manual annotations. Each filament fragment was manually annotated at its selected best-focus plane, although the same structure could remain visible across neighboring optical sections. Consequently, requiring exact slice-level agreement could underestimate spatial correspondence. To address this limitation, an axial-tolerant overlap rate (ATOR) was defined by first computing the maximum overlap fraction of each focused OBB within a specified axial neighborhood,

$$c_i(k) = \max_{|z-z_i|\le k} \frac{|B_i \cap G_z|}{B_i} \tag{10}$$

where B_i denotes the rasterized region of the i-th focused OBB, z_i is its corresponding optical section, and k defines the maximum allowed axial displacement in number of slices. The overlap fraction was calculated relative to the full rasterized area of each OBB. The overall ATOR was then calculated as the percentage of focused OBBs satisfying $c_i(k)$,

$$ATOR(k,\delta) = \frac{100}{N}\sum_{i=1}^{N} 1[c_i(k) \ge \delta] \tag{11}$$

where N denotes the total number of focused OBBs, δ represents the minimum required overlap fraction, and $1[\cdot]$ is an indicator function equal to one when the condition is satisfied and zero otherwise. ATOR was evaluated using axial tolerances from 0 to ±35 μm in increments of 5 μm and minimum overlap criteria of 10%, 20%, and 30%. These parameters were examined through sensitivity analysis rather than selected as optimal thresholds.

To assess whether the observed correspondence exceeded that obtained under spatial displacement, a translation-based randomization control was performed. For each trial r, all focused OBBs were displaced by a common random vector in the XY plane, where the displacement magnitude was sampled uniformly within a predefined range and its direction uniformly over 0–2π. The relative arrangement, dimensions, orientations, and axial positions of the OBBs were preserved, while the manual segmentation remained fixed. Three displacement ranges were evaluated (25–50, 50–100, and 100–200 μm), with 100 trials per range. For each trial, ATOR was recalculated across all axial tolerances and overlap criteria. Statistical significance was assessed using a one-sided empirical Monte Carlo test (63),

$$p_{emp}(k,\delta) = \frac{1 + \sum_{r=1}^{R} 1\left[ATOR^{(r)}(k,\delta) \ge ATOR_{obs}(k,\delta)\right]}{R+1} \tag{12}$$

where R denotes the number of randomized trials, $ATOR^{(r)}$ is the overlap rate obtained in trial (r), and $ATOR_{obs}(k, \delta)$ is the overlap rate of the original predictions. This analysis evaluates spatial correspondence relative to the specified randomization protocol.

4. Results

4.1 Dataset characterization and annotation statistics

Following the annotation procedure described in *Section 3.3*, the final OBB dataset comprised 21 annotated Z-stacks containing 1,205 slice-level image–label pairs acquired at the native image resolution of 2928 × 2928 pixels. The number of annotated slices per stack ranged from 20 to 193, reflecting differences in the depth and spatial extent of the imaged mycelial architectures. Stacks 0–8 were acquired using a ×20 objective, whereas stacks 9–20 were acquired using a ×4 objective, exposing the detector to filament fragments across multiple spatial scales.

After subdivision into mosaic-level samples and exclusion of tiles containing no annotated filament fragments, the final training dataset comprised 7,657 annotated 732 × 732-pixel mosaics. The number of retained mosaics varied among stacks according to the spatial extent and density of the fungal network. A detailed summary of the annotated slices, retained mosaics, and acquisition objective for each Z-stack is presented in Table 3. Overall, the resulting dataset spans multiple imaging scales and a broad range of filament densities and structural complexities, providing a diverse collection of annotated filament fragments for training and evaluating the OBB detection model.

Table 3. Per-stack summary of annotated images, mosaics, and acquisition magnifications.

Objective	Stack ID	Annotated slices	Annotated mosaics	Stack ID	Annotated slices	Annotated mosaics

		(2928 x 2928)	(732 x 732)		(2928 x 2928)	(732 x 732)
× 20	stack000	51	453	stack005	35	262
	stack001	20	183	stack006	31	252
	stack002	22	132	stack007	58	510
	stack003	26	182	stack008	49	306
	stack004	78	499			
× 4	stack009	30	225	stack015	25	125
	stack010	30	267	stack016	63	639
	stack011	40	293	stack017	40	280
	stack012	101	580	stack018	193	1031
	stack013	106	627	stack019	56	168
	stack014	101	303	stack020	50	340
				Total	**1205**	**7657**

4.2 Performance of OBB-based filament detection

Detection performance was evaluated using five-fold cross-validation at the mosaic level. The folds were approximately balanced, containing between 1,472 and 1,560 annotated mosaics (approximately 20% of the dataset per fold) with each validation fold including mosaics originating from 15–19 Z-stacks and spanning both acquisition objectives. Training and validation losses for box regression, classification, and distribution focal loss converged consistently across folds (Additional file 1: Fig. S1). Validation performance consistently stabilized between epochs 70 and 80 across all folds, without evidence of substantial overfitting. As model checkpoints were stored every five epochs, epoch 75 was selected as the operating model for all subsequent analyses.

The operating confidence and IoU thresholds were selected by maximizing the composite score S across the validation folds. As shown in Fig. 3, the optimal operating point was obtained at an IoU threshold of 0.60 and a confidence threshold of 0.126. The broad plateau surrounding the selected operating point indicates that detector performance is

robust to small variations in the inference thresholds, particularly with respect to the IoU threshold. These thresholds were therefore retained for all subsequent experiments.

At the selected operating point, the detector achieved an instance-level F1-score of 0.667 ± 0.024 and a coverage-based F1-score of 0.779 ± 0.024 (Table 4). The relatively small standard deviations across the five validation folds indicate consistent detection performance despite variations in biological samples and imaging conditions. The corresponding composite score of $S = 0.723 \pm 0.019$ further demonstrates robust overall detection performance.

Figure 3. Selection of the operating inference thresholds using five-fold cross-validation. The composite score S is shown as a function of (a) the IoU threshold and (b) the confidence threshold. Colored curves correspond to individual validation folds, the solid black curve denotes the mean performance, and the shaded region indicates ±1 standard deviation. Red markers identify the selected operating thresholds (IoU = 0.60, confidence = 0.126), corresponding to the maximum mean composite score across the five validation folds.

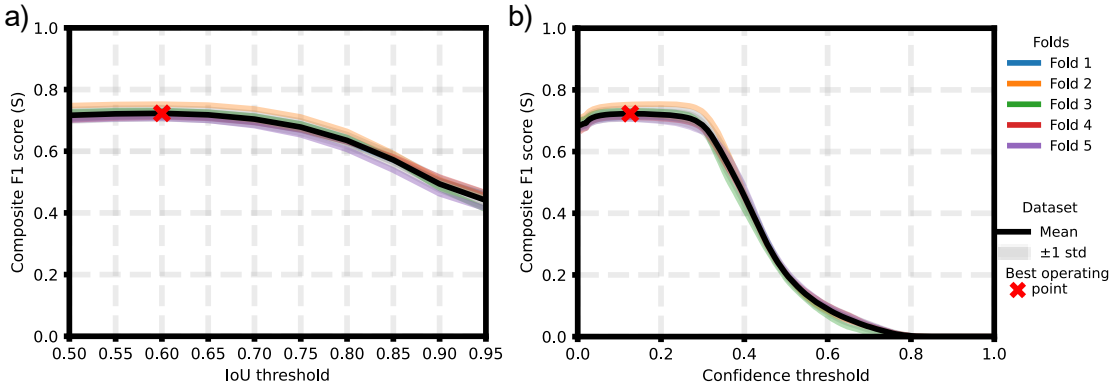


Table 4. Detection performance obtained by five-fold cross-validation at the selected operating point (IoU = 0.60, confidence = 0.126). Results are reported as mean ± standard deviation across

672 the validation folds. Detection metrics correspond to instance-level matching, coverage metrics
673 quantify spatial agreement, and the composite score S summarizes both evaluation schemes.

Metric	Precision	Recall	F1-Score	mAP50
Detection	0.647 ± 0.015	0.690 ± 0.042	0.667 ± 0.024	0.603±0.030
Coverage	0.839 ± 0.023	0.729 ± 0.055	0.779 ± 0.024	
Composite score (S)	-	-	0.723 ± 0.019	

674 Following cross-validation, a final detector was trained using the complete annotated
675 dataset, retaining the selected operating epoch (epoch 75) and inference thresholds (IoU =
676 0.60, confidence = 0.126). The final model was subsequently evaluated across all 21
677 annotated Z-stacks to characterize its operational behavior. Because these data were also used
678 for training, the reported values do not constitute an independent estimate of generalization.
679 Instead, they characterize the performance of the final detector under the conditions
680 represented in the annotated dataset. The model achieved slice-level mean F1-scores of 0.720
681 ± 0.080 and 0.808 ± 0.095 for the detection- and coverage-based evaluations, respectively,
682 yielding a composite score of 0.764 ± 0.072 (Table 5).
683 Runtime measurements were obtained on a workstation equipped with a 13th Gen Intel
684 Core i9-13900KF processor, 128 GB RAM, and an NVIDIA RTX A5000 GPU with 24 GB
685 VRAM. The average computation time was 70.3 ± 8.2 ms per slice (14.4 ± 1.6 slices/s),
686 increasing to 180.9 ± 20.1 ms per slice (5.6 ± 0.6 slices/s) for end-to-end processing.

687 **Table 5.** Slice-level detection performance and runtime statistics of the final model. Detection
688 and coverage metrics are reported as mean ± standard deviation, with ranges in parentheses.
689 Runtime statistics include both model computation time and end-to-end processing time

(*Compute* includes preprocessing, model inference, and post-processing, and *end-to-end* additionally accounts for data loading and result writing).

Metric	Value
P_{det}	0.651 ± 0.087 (0.138-0.917)
R_{det}	0.819 ± 0.109 (0.098-1.000)
$F1_{det}$	0.720 ± 0.080 (0.115-0.880)
P_{cov}	0.835 ± 0.149 (0.130-0.995)
R_{cov}	0.803 ± 0.058 (0.396-0.943)
$F1_{cov}$	0.808 ± 0.095 (0.227-0.929)
S	0.764 ± 0.072 (0.349-0.895)
Detections / slice	117.080 ± 77.413 (9.000–403.000)
Compute (ms)	70.334 ± 8.233 (52.325–109.643)
End-to-End (ms)	180.862 ± 20.068 (122.461–314.413)
Compute (slices/s)	14.406 ± 1.631 (9.121–19.111)
End-to-End (slices/s)	5.597 ± 0.627 (3.181-8.166)

Substantial variability in the number of detections was observed across slices (117.1 ± 77.4 detections per slice), reflecting differences in filament density, morphology, and structural complexity throughout the dataset (Table 5). Despite this heterogeneity, the slice-level distributions of the detection, coverage, and composite metrics remain concentrated around their central values (Fig. 4), although noticeable variability remained across individual slices. A systematic difference is nevertheless observed between the two evaluation schemes, with detection exhibiting lower precision but higher recall than coverage. Consequently, the composite score S lies between the corresponding detection- and coverage-based F1-scores, providing a balanced summary of both instance-level localization and spatial agreement. The origin of these systematic differences is examined through representative examples in Fig. 5.

Figure 4. Distribution of slice-level performance metrics for the final detector across all 1,205 annotated image slices. Violin plots summarize the distributions of detection and coverage precision, recall, and F1-score, together with the composite score S. Dashed lines indicate the median and interquartile range (Q1–Q3).

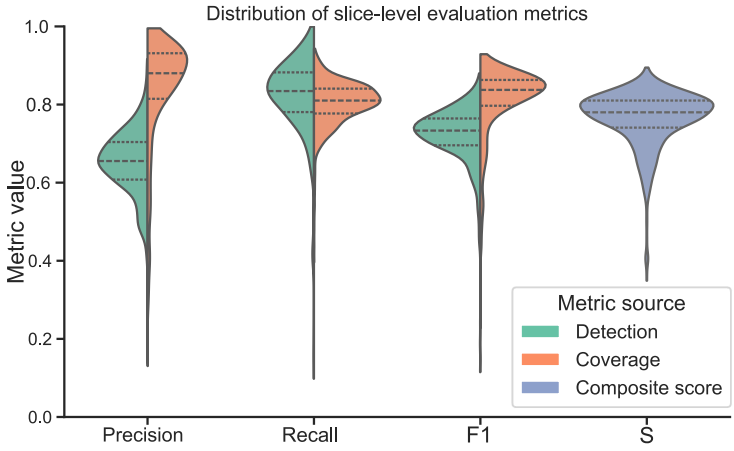
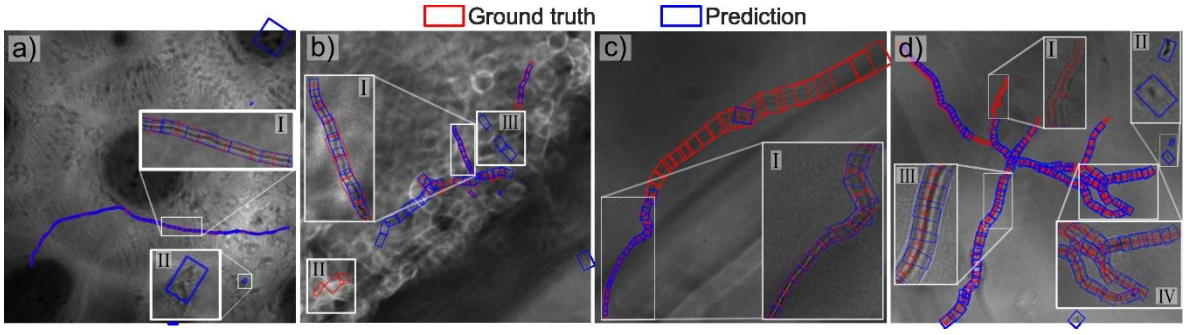


Figure 5. Representative examples of detection results across different slices. Ground-truth annotations are shown in red and model predictions in blue. Insets (I–IV) highlight local regions illustrating typical agreement and discrepancies between annotations and predictions, including curved filaments, overlapping detections, and small isolated structures.



The observed reduction in detection precision arises from the interaction between the continuous nature of filamentous structures and the discrete constraints imposed by instance-level evaluation. Filaments are represented as sequences of local OBBs, and multiple

714 geometrically valid configurations may describe the same underlying structure. However,
715 evaluation is based on a one-to-one assignment between predictions and ground-truth
716 annotations, such that only a single prediction can be matched to each annotated instance. As
717 a result, additional predictions that correctly cover the filament are counted as false positives
718 (FP), even when their overlap exceeds the IoU threshold (Fig. 5bI and 5cI). This effect is
719 particularly evident in curved filaments, where several valid orientations and placements
720 exist along the same structure (Fig. 5cI), leading to cases in which the number of predictions
721 exceeds the number of annotations while still preserving correct spatial coverage.

722 Additional discrepancies arise from the geometric sensitivity of instance-level
723 evaluation. Predictions that are spatially well localized but differ slightly in orientation or
724 extent may fail to satisfy the IoU criterion, producing both an unmatched prediction (FP) and
725 an unmatched annotation (FN) (Fig. 5dIII,dIV). Such differences frequently occur in regions
726 affected by halo artifacts or local intensity variations that alter the apparent filament width.
727 Other false positives originate from stain-like structures resembling filament fragments (Fig.
728 5aII,dII) or from plausible filament segments omitted during manual annotation (Fig. 5bIII),
729 particularly in low-contrast or partially defocused regions. Residual FNs are primarily
730 associated with weakly visible filaments, structures spanning multiple focal planes, or
731 detections removed by the confidence threshold (Fig. 5bII,c,dI).

732 In contrast, coverage-based evaluation assesses spatial agreement rather than one-to-one
733 correspondence between instances. Consequently, duplicated detections along the same
734 filament have only a minor influence on precision because they largely overlap the same
735 spatial region. However, coverage metrics become more sensitive when FP detections occupy

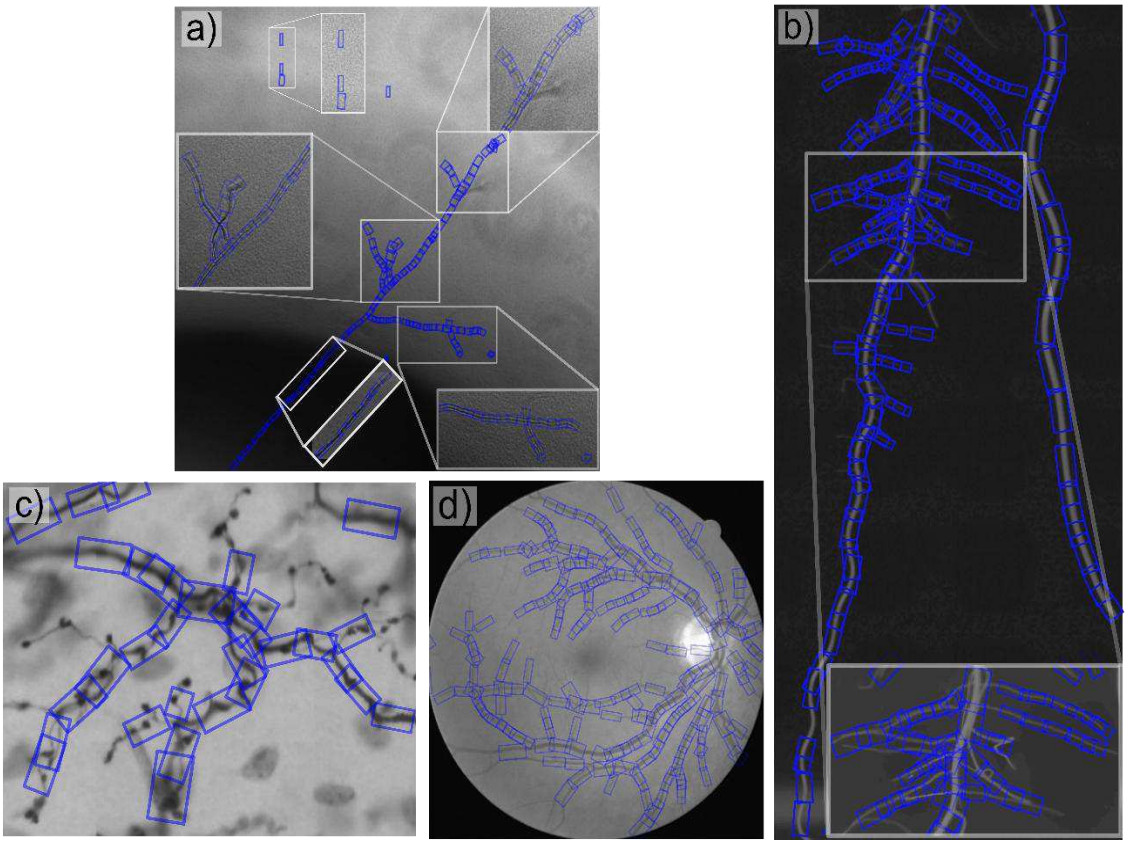
736 relatively large areas. This occurs, for example, for isolated stain-like structures or large
737 spurious OBBs (Fig. 5aII,dII), where the predicted support region increases despite the
738 absence of a corresponding filament. Such detections generate a proportionally larger FP area
739 than they do under instance-level evaluation. Nevertheless, these cases are relatively
740 infrequent, as reflected by the consistently higher precision obtained with coverage-based
741 evaluation compared with detection (Fig. 4 and Table 5).

742 Neither evaluation paradigm alone provides a complete assessment of filament detection
743 performance. Instance-level detection metrics penalize redundant yet geometrically valid
744 predictions through one-to-one matching, potentially underestimating performance for
745 continuous filamentous structures. Conversely, coverage-based evaluation may achieve high
746 scores despite fragmented or disconnected detections, provided sufficient spatial overlap is
747 preserved. Consequently, relying on either metric in isolation may lead to misleading
748 conclusions regarding detection performance. The proposed composite score (S) addresses
749 this limitation by integrating instance-level localization and spatial coverage into a single
750 measure, providing a balanced summary of both evaluation perspectives.

751 Figure 6 presents qualitative results on image types not included during training to assess
752 the transferability of the proposed detector. In Fig. 6a, the model is applied to mycelium
753 images acquired at an intermediate magnification (×10), whereas training was performed
754 exclusively using ×4 and ×20 images. Figures 6b–d further evaluate the detector on plant
755 roots, fiber-like cells, and retinal vessels. Across all cases, the predicted OBBs consistently
756 follow elongated and branching structures, preserving their local orientation and continuity
757 despite differences in imaging modality, scale, and biological domain. Most discrepancies

occur in densely connected regions or thin, low-contrast structures, where redundant detections or isolated FPs are occasionally observed. Although qualitative, these results suggest that the learned representation captures general geometric properties of filamentous networks and may be transferable to related biological and biomedical imaging applications.

Figure 6. Qualitative detection results in unseen image datasets. Predicted OBBs (blue) are overlaid on the input images. (a) Mycelium (×10 objective). (b) Plant roots. (c) Fiber-like cells. (d) Retinal vessels. Insets show representative local detections.



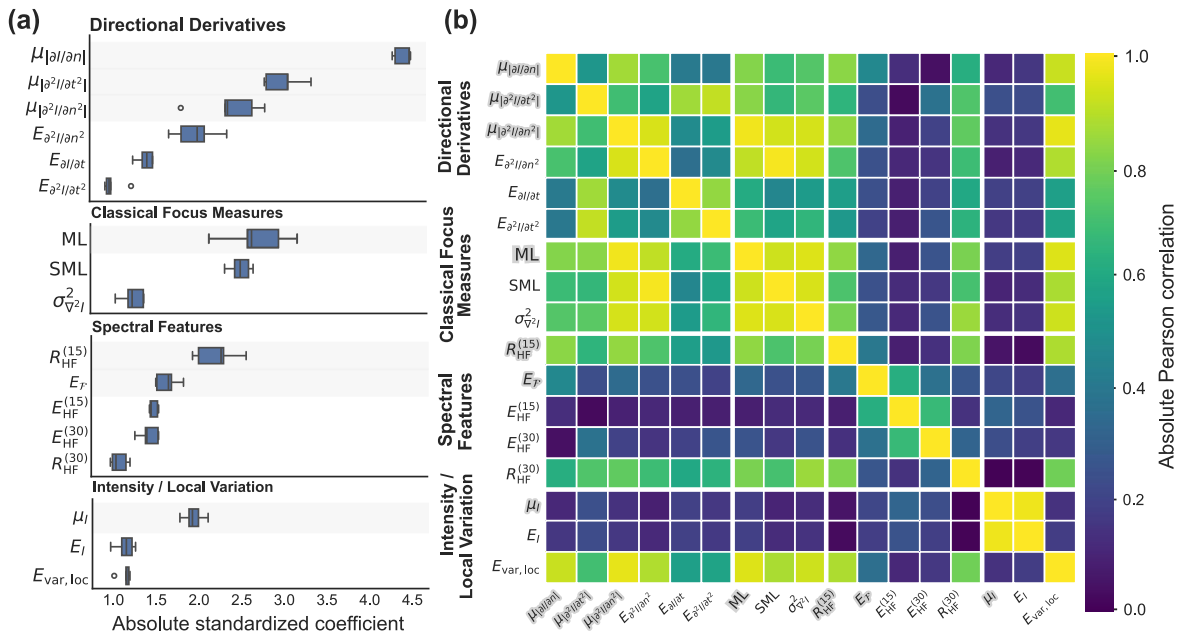
4.3 Focus classification and depth consistency

The relevance of the proposed features was first assessed through the distribution of logistic regression coefficients across cross-validation folds (Fig. 7a). Features with substantial contributions to the logistic-regression model were distributed across multiple descriptor families. In particular, the directional derivative features $\mu_{|\partial I/\partial n|}$, $\mu_{|\partial^2 I/\partial t^2|}$, and

771 $E_{\partial^2 I/\partial n^2}$ exhibit consistently high coefficients, comparable to or exceeding those of classical
772 focus measures such as SML, and $\sigma^2_{\nabla^2 I}$. These results suggest that orientation-aware
773 derivatives contain substantial focus-related information, with contributions comparable to
774 those of several classical focus measures.

775 The correlation matrix (Fig. 7b) revealed a clear block structure consistent with the
776 feature categories. Several directional derivatives and classical focus measures exhibited
777 strong intra- and inter-group correlations, indicating substantial overlap in the local edge-
778 and sharpness-related information captured by these descriptors. In contrast, the spectral
779 descriptors showed lower correlations with the directional features, particularly (E_F), while
780 the intensity descriptor (μ_I) remained weakly correlated with all other features outside the
781 intensity group. These observations indicate that combining directional, spectral, and
782 intensity descriptors is more likely to provide complementary information than combining
783 multiple highly correlated features from the same category.

784 **Figure 7.** Feature relevance and correlation analysis: a) Absolute logistic regression coefficients
785 across folds for all features, grouped by category (Table 2). Gray bands highlight the subset of
786 features selected as the most relevant within each category for subsequent triplet construction.
787 b) Pairwise Pearson correlation matrix showing intra- and inter-group relationships. Gray-
788 highlighted rows and columns correspond to the same subset of selected features identified in a).



789

790 Based on these observations, six cross-category feature triplets were constructed by
791 combining representative directional, spectral, and intensity descriptors. In addition, a control
792 triplet composed of highly ranked but correlated descriptors was evaluated to assess the effect
793 of feature redundancy. Among the evaluated representations (Table 6), the triplet combining
794 $\mu_{|\partial^2 I/\partial n^2|}$, $R_{HF}^{(15)}$, and $\mu_I$achieved the highest validation ROC-AUC (0.952) and PR-AUC
795 (0.945), indicating the strongest threshold-independent discrimination. In contrast, the
796 combination of $\mu_{|\partial I/\partial n|}$, $R_{HF}^{(15)}$, and μ_I achieved the highest validation F1-score (0.851),
797 with a precision of 0.749 and a recall of 0.986.

798 Based on its validation F1-score, the triplet $\mu_{|\partial I/\partial n|} + R_{HF}^{(15)} + \mu_I$was selected as the final
799 focus representation. Its combination of directional, spectral, and intensity information
800 provided the best balance between precision and recall among the evaluated candidates. In
801 contrast, the control triplet $(\mu_{|\partial I/\partial n|}, \mu_{|\partial^2 I/\partial t^2|}, \mathrm{ML})$, composed of individually relevant but
802 correlated descriptors achieved a lower validation F1-score of 0.721, supporting the benefit
803 of combining complementary sources of focus-related information while limiting feature
804 redundancy.

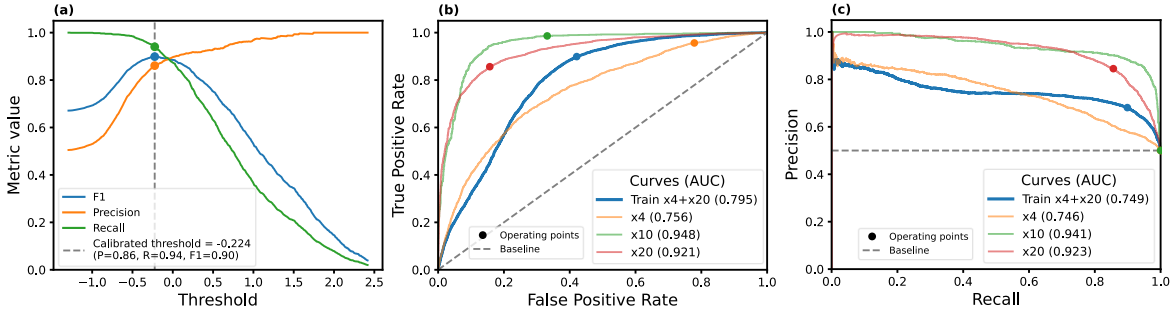
Table 6. Performance of focus descriptor triplets evaluated using P, R, F1, ROC-AUC, and PR-AUC. The best values for each metric are shown in bold. The triplet selected for subsequent analysis is indicated by bold text in the first column.

	Train					Valid				
	P	R	F1	ROC	PR	P	R	F1	ROC	PR
$\mu_{\lvert\partial^2 I/\partial n^2\rvert}, R_{HF}^{(15)}, \mu_I$	0.654	0.895	0.755	0.773	0.744	0.697	0.992	0.819	**0.952**	**0.945**
$\boldsymbol{\mu_{\lvert\partial I/\partial n\rvert}, R_{HF}^{(15)}, \mu_I}$	**0.681**	0.899	**0.775**	0.795	0.749	**0.749**	0.986	**0.851**	0.948	0.941
$\mu_{\lvert\partial^2 I/\partial t^2\rvert}, R_{HF}^{(15)}, \mu_I$	0.593	0.915	0.720	0.704	0.678	0.576	0.997	0.730	0.934	0.925
$\mu_{\lvert\partial I/\partial n\rvert}, E_F, \mu_I$	0.670	0.838	0.745	**0.807**	**0.796**	0.701	0.928	0.798	0.892	0.889
$\mu_{\lvert\partial^2 I/\partial n^2\rvert}, E_F, \mu_I$	0.651	0.844	0.735	0.789	0.773	0.647	0.922	0.76	0.856	0.848
$\mu_{\lvert\partial^2 I/\partial t^2\rvert}, E_F, \mu_I$	0.605	**0.953**	0.740	0.671	0.625	0.562	**1.000**	0.719	0.833	0.750
$\mu_{\lvert\partial I/\partial n\rvert}, \mu_{\lvert\partial^2 I/\partial t^2\rvert}, \mathrm{ML}$	0.599	0.936	0.730	0.692	0.674	0.564	0.999	0.721	0.888	0.845

For the selected triplet, the two-component PCA representation captured 94.74% of the variance in the standardized feature space, with PC1 and PC2 explaining 60.94% and 33.80%, respectively. As shown in Table 7, PC1 exhibited comparable positive loadings for the normal gradient (0.709) and high-frequency ratio (0.704), with a substantially smaller contribution from mean intensity (0.050). In contrast, PC2 was predominantly associated with mean intensity, with a loading of 0.989.

The selected PCA-based focus representation was subsequently applied to the ×10 validation data without refitting. Threshold sensitivity analysis identified a calibrated operating threshold of $\tau = -0.224$, yielding an F1-score of 0.90, with a precision of 0.86 and a recall of 0.94 (Fig. 8a). The ROC and precision–recall curves further show that the continuous focus score maintained discriminative performance across acquisition objectives (Fig. 8b,c), with ROC-AUC values of 0.756, 0.948, and 0.921 and PR-AUC values of 0.746, 0.941, and 0.923 for the ×4, ×10, and ×20 datasets, respectively.

Figure 8. Performance and operating-threshold calibration of the selected PCA-based focus representation. (a) Precision, recall, and F1-score as a function of the decision threshold on the ×10 validation dataset. The dashed vertical line indicates the calibrated operating threshold ($\tau = -0.224$) selected by maximizing the F1-score. (b) Receiver operating characteristic (ROC) curves and (c) precision–recall (PR) curves for the combined ×4–×20 model-development data and the objective-specific datasets. Markers indicate the operating points obtained using the calibrated threshold.



The continuous focus representation showed objective-dependent discriminative performance (Fig. 8b,c). The ×4 dataset achieved lower ROC-AUC (0.756) and PR-AUC (0.746) than the ×20 (0.921 and 0.923, respectively) and ×10 (0.948 and 0.941, respectively) datasets. The lower performance observed at ×4 is consistent with the reduced spatial detail available at lower magnification, where filament-associated OBB patches are represented by fewer pixels and fine intensity variations may be less pronounced. In contrast, the higher and comparable performance obtained for the ×10 and ×20 datasets indicates that the selected focus representation remains discriminative across these imaging magnifications. The inclusion of the more challenging ×4 data in the model-development dataset also contributed to lower aggregate development performance relative to the ×10 validation results.

Table 7. Parameters of the selected PCA-based focus model

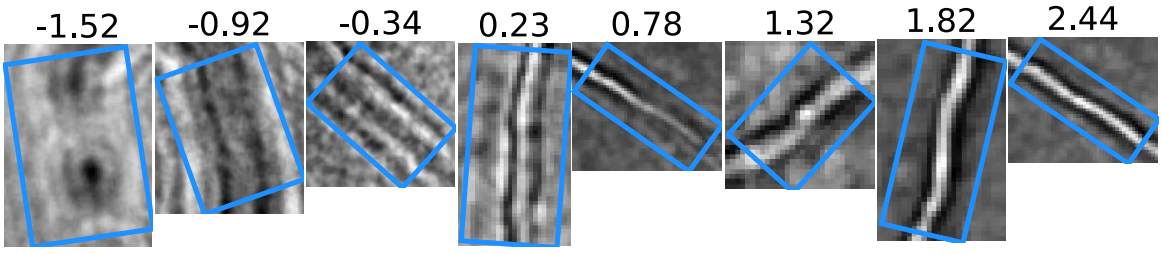
Component	Symbol	Value		
Features	x	$[\mu_{\lvert\partial I/\partial n\rvert},\ R_{HF}^{(15)},\mu_I]$		
Standardization	$\mu_k \pm \sigma_k$	0.564 ± 0.300	0.064 ± 0.059	0.447 ± 0.093
	PC1	[0.709,	0.704,	0.050]

PCA components	PC2	[0.061, −0.132, 0.989]
Explained variance ratio	λ_k	0.609,0.338
PC weights used in score	ω_k	0.643,0.357
PC orientation signs	$sign_k$	[1 −1]
Score statistics	$\mu_s \pm \sigma_s$	0.000 ± 0.941
Threshold	τ	-0.224

The parameters of the final focus model are summarized in Table 7. Feature-standardization statistics, PCA components, explained-variance ratios, component weights, orientation signs, and score-standardization statistics were estimated from the ×4 and ×20 model-development data. The final operating threshold was subsequently calibrated on the ×10 validation dataset, yielding $\tau = -0.224$.

The distribution of the normalized focus scores further illustrates the relationship between the continuous focus representation and the annotated focus classes (Fig. 9). Defocused samples were predominantly associated with lower score values, whereas focused samples were shifted toward higher values. The partial overlap between the two distributions reflects the gradual variation in filament sharpness across adjacent optical sections, consistent with modeling focus as a continuous property of filament appearance. The calibrated threshold $\tau = -0.224$ provides the operating boundary used to convert this continuous representation into the binary focus labels required for depth-aware filtering.

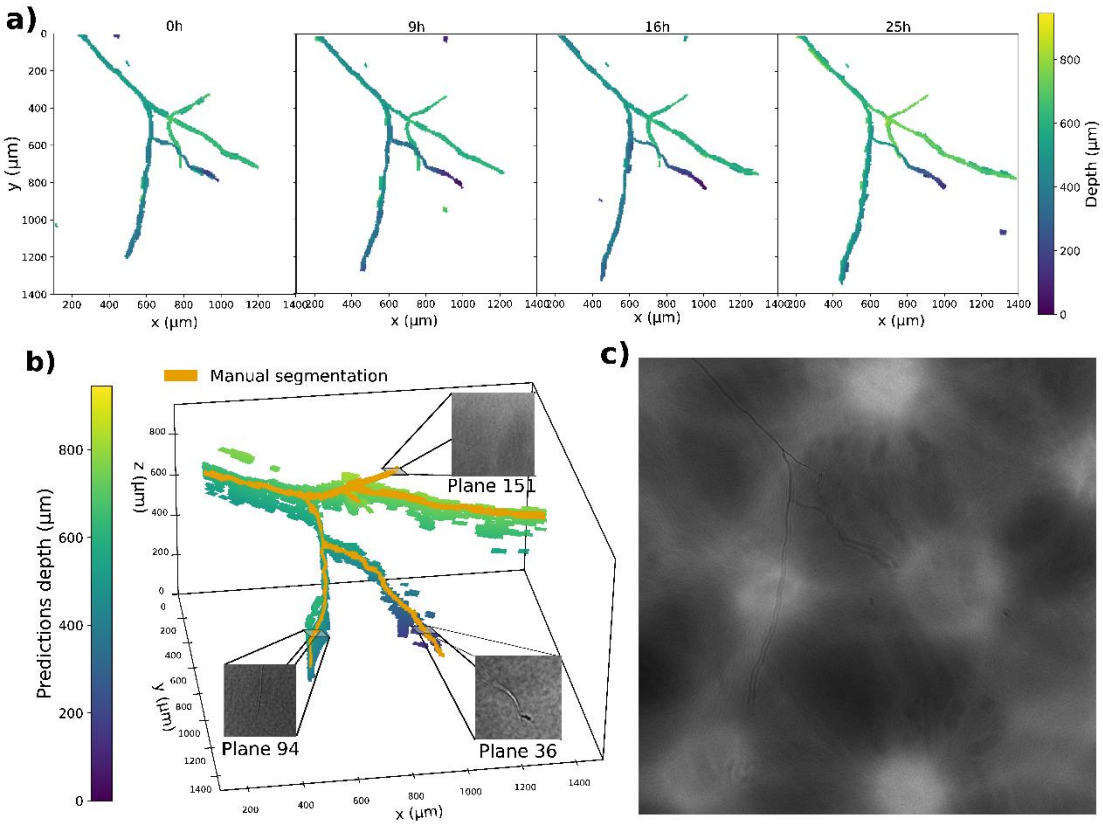
Figure 9. OBB patches ordered by increasing focus score for the triplet $\mu_{|\partial I/\partial n|}$, $R_{HF}^{(15)}$, μ_I. Lower scores correspond to out-of-focus regions, while higher scores indicate sharper filament structures.



4.4 Three-dimensional reconstruction and temporal evolution

The proposed pipeline was tested on a separate 25-hour time-lapse experiment comprising 48 Z-stacks acquired at approximately 30-min intervals using a ×10 objective. The time-lapse data were not used for OBB detector training. As described in *Section 3.5*, the first time point was used to evaluate the fixed PCA-based focus representation and calibrate the ×10 operating threshold. Representative reconstructions at 0, 9, 16, and 25 h are shown in Fig. 10a, illustrating the temporal evolution of the reconstructed filamentous structure. The complete time-lapse reconstruction is provided in Additional file 2: Video S1, while an additional application to a more structurally complex fungal network is shown in Additional file 1: Fig. S2.

Figure 10. Three-dimensional reconstruction and temporal evolution of filament structures. (a) OBB-based three-dimensional reconstructions at representative time points (0, 9, 16, and 25 h), illustrating the temporal evolution of the filamentous structure. Colors indicate reconstructed depth along the z-axis. (b) Qualitative comparison between the OBB-based three-dimensional reconstruction (color-coded) and a manual segmentation of the complete filamentous structure (orange) at the final time point. Insets show representative microscopy images from planes 36, 94, and 151, illustrating portions of the structure located at distinct axial positions within the Z-stack. (c) Raw bright-field microscopy image from plane 151 of the final Z-stack, showing the original image data corresponding to the reconstructed structure.



876

877 The reconstructed OBBs captured the temporal evolution of the fungal structure while
878 preserving its three-dimensional spatial organization across successive Z-stacks (Fig. 10a).
879 At the final time point, the reconstruction showed spatial trajectories consistent with the
880 manual segmentation (Fig. 10b), while the depth-coded representation revealed the axial
881 distribution of filament fragments. Focus-aware filtering reduced defocused detections,
882 yielding a coherent volumetric representation that facilitates the distinction between filament
883 crossings and branching structures. Residual discrepancies were mainly observed in low-
884 contrast regions and areas of uneven illumination, as illustrated by the raw microscopy image
885 (Fig. 10c). Although explicit connections between neighboring OBBs were not established,
886 incorporating local spatial and geometric context in a future connectivity analysis could
887 enable more selective removal of false-positive detections while preserving structurally
888 consistent filament fragments.

Table 8. ATOR and spatial randomization controls. Original ATOR values are compared with the mean ± standard deviation obtained from 100 random XY translations of 25–50 μm. Results are reported for axial tolerances of 0–35 μm and minimum OBB overlap criteria of 10%, 20%, and 30%.

	$ATOR \geq 10\%$		$ATOR \geq 20\%$		$ATOR \geq 30\%$	
Axial tolerance (um)	Original (%)	Control mean ± SD (%)	Original (%)	Control mean ± SD (%)	Original (%)	Control mean ± SD (%)
0	30.35	4.35 ± 3.12	27.70	3.15 ± 2.40	24.63	2.12 ± 1.82
5	40.30	6.44 ± 4.34	38.27	4.91 ± 3.48	35.20	3.45 ± 2.76
10	48.40	8.20 ± 5.46	46.33	6.39 ± 4.47	43.64	4.54 ± 3.58
15	56.46	9.79 ± 6.52	54.29	7.70 ± 5.38	51.15	5.52 ± 4.36
20	63.80	11.28 ± 7.52	61.56	8.96 ± 6.27	58.15	6.41 ± 5.10
25	69.79	12.63 ± 8.40	67.55	10.07 ± 7.03	64.17	7.20 ± 5.72
30	74.06	13.82 ± 9.10	72.24	11.00 ± 7.67	68.65	7.86 ± 6.23
35	77.23	14.77 ± 9.64	75.75	11.75 ± 8.14	71.96	8.39 ± 6.62

Quantitative evaluation of the reconstructed filament distribution yielded a global pixel-wise recall of 0.9284, indicating substantial coverage of the manual annotations by focused OBB predictions. The ATOR further assessed the spatial correspondence of individual predictions across neighboring optical sections (Table 8). At zero axial tolerance, ATOR ranged from 24.63% to 30.35% across the three overlap criteria, increasing to 71.96–77.23% at ±35 μm. At ±35 μm and a minimum overlap threshold of 20%, the original reconstruction achieved an ATOR of 75.75%, substantially exceeding the randomized control with XY displacements of 25–50 μm (11.75 ± 8.14%), a difference of 64.00 percentage points. None of the 100 randomized trials reached the observed value (one-sided empirical Monte Carlo test, $p = 0.0099$). Larger displacement ranges of 50–100 μm and 100–200 μm yielded even lower overlap rates of 6.03 ± 4.26% and 2.59 ± 2.35%, respectively (Additional file 1: Tables S2, S3). These results demonstrate that the reconstructed filament distribution exhibits substantially greater spatial correspondence with manual annotations than the randomly

displaced controls, although explicit topological connectivity between neighboring OBBs remains unresolved.

4. Discussion

The reconstruction of filamentous structures from microscopy images presents challenges that are not adequately addressed by conventional object detection or segmentation frameworks. Unlike compact objects, filamentous structures are continuous, anisotropic, and frequently exhibit low contrast, overlap, and poorly defined boundaries. In this work, these challenges are addressed by representing local filament fragments as OBBs, rather than binary masks or axis-aligned bounding boxes. By explicitly encoding position, orientation, and scale, OBBs provide a compact geometric representation that is naturally adapted to elongated structures and facilitates downstream processing, including focus estimation and three-dimensional reconstruction. Although demonstrated on fungal structures, the proposed representation is not specific to this application and may provide a useful representation for other filamentous biological systems, including vascular, neuronal, and plant root networks.

Unlike conventional object detection tasks, filamentous structures admit multiple valid OBB decompositions, complicating the interpretation of standard detection metrics. Although oriented object detection has been extensively studied across diverse applications (64), differences in object geometry, image characteristics, and annotation strategies limit direct performance comparisons. Nevertheless, to contextualize the obtained mAP50 of 60.30%, previous studies have reported values above 90% on HRSC2016 (65) and approximately 35–45% on the more challenging FAIR1M dataset (66). These comparisons illustrate the variability of detection performance across domains rather than establish a direct benchmark for fungal filaments. Consequently, complementary instance-level and spatial-coverage evaluations provide a more comprehensive characterization of filament detection performance.

A second contribution of this work concerns the formulation of focus estimation as a continuous, data-guided focus representation. Classical SFF methods rely on selecting a predefined focus operator and evaluation window, both of which strongly influence reconstruction quality and require careful parameter tuning. In contrast, the proposed framework replaces fixed windows with OBB-aligned patches that adapt to the geometry and orientation of each detected filament fragment. Rather than relying on a single focus operator, complementary directional, spectral, and intensity descriptors are integrated into a continuous focus score using PCA. This formulation models focus as a gradual property of filament appearance rather than a binary state, reliance on a single focus operator and providing a geometry-aware focus representation that transferred across acquisition objectives while allowing objective-specific calibration of the operating threshold. This distinction differentiates the proposed approach from previous microscopy-based reconstruction frameworks based on conventional SFF formulations.

Despite these encouraging results, several limitations remain. The proposed framework relies on manually annotated OBBs and focus labels, making dataset construction labor-intensive and introducing a degree of subjectivity, particularly near focal transitions where binary focus labels may be ambiguous. Evaluation also remains challenging because conventional object-detection and segmentation metrics do not fully capture the characteristics of continuous filamentous structures, for which multiple geometrically plausible OBB representations may exist. The coverage-based metric complements one-to-one instance matching by quantifying spatial agreement, while the composite score integrates both perspectives; however, neither directly measures structural or topological reconstruction accuracy, which should be addressed in future evaluation strategies.

The quantitative spatial evaluation provided additional evidence of correspondence between the reconstructed filament distribution and manual annotations. The increase in ATOR with axial tolerance is consistent with the annotation procedure, in which filament

fragments were labeled at their selected best-focus planes despite remaining visible across neighboring optical sections. Furthermore, the substantially higher overlap rates observed for the original reconstruction compared with spatially randomized controls support the spatial specificity of the predictions. Nevertheless, ATOR evaluates correspondence within a predefined axial neighborhood rather than absolute depth accuracy, and the evaluation was limited to a single Z-stack. The absence of independent three-dimensional ground truth and explicit connectivity between neighboring OBBs therefore remains a limitation for evaluating geometric and topological reconstruction accuracy.

Establishing explicit filament connectivity would substantially extend the biological applicability of the framework. Once reconstructed structures are connected into three-dimensional graphs, and temporal correspondence is established across successive time points, they could support the identification of branches and tips, characterization of path geometry, and quantitative analysis of changes in network architecture over time. Such representations could enable measurements of filament growth, branching dynamics, and structural responses to environmental conditions, providing a basis for studying fungal organization in both space and time.

6. Conclusions

This work presents a data-driven framework based on YOLO-OBB detection for the three-dimensional reconstruction of filamentous structures from bright-field microscopy. Developed and evaluated using the arbuscular mycorrhizal fungus Rhizophagus irregularis, the detector also demonstrated qualitative applicability to other biological images containing similar filamentous morphologies. By integrating orientation-aware detection with continuous focus modeling for depth estimation, the framework enabled three-dimensional reconstruction of fungal structures throughout a 25-hour time-lapse experiment. Quantitative evaluation yielded a global pixel-wise recall of 0.9284 and an axial-tolerant overlap rate

(ATOR) of 75.75% at ±35 μm with a minimum overlap criterion of 20%, substantially exceeding the spatially randomized controls. These findings support the potential of OBB-based representations for automated three-dimensional analysis of filamentous biological structures. Although explicit topological connectivity remains unresolved, the proposed framework provides a foundation for future quantitative investigations of fungal network architecture and growth dynamics.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Data availability statement

The source code, trained models, configuration files, and reproducibility scripts are available in the `paper-release` branch of the GitLab repository: https://gitlab.iit.it/dalia.dominguez/mycelium_analysis/-/tree/paper-release. The microscopy dataset used in this study is publicly available at Zenodo (https://doi.org/10.5281/zenodo.23135747).

Competing interests

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.

Funding

This work received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Program (Grant Agreement No. 101003304, I-Wood).

CRediT authorship contribution statement

D.Y.D.-J.: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. M.R.: Writing – review & editing, Methodology, Investigation, Data curation. E.D.D.: Writing – review & editing, Supervision, Methodology, Conceptualization. B.M.: Writing – review & editing, Supervision, Resources, Funding acquisition, Conceptualization. All authors reviewed and approved the final manuscript.

Acknowledgements

Dalia Y. Domínguez-Jiménez acknowledges financial support from SECIHTI, Mexico, through the 2025 Postgraduate Studies in Science and Humanities Abroad grant.

Additional files

Additional file 1

File name: Supplementary_Material.docx

File format: DOCX

Title: Supplementary figures, tables, and methodological details.

Description: Contains YOLO11n-OBB training and validation loss curves (Figure S1), mathematical definitions of the 18 focus descriptors (Table S1), an additional reconstruction example of a complex fungal network (Figure S2), and quantitative evaluations of the 3D reconstruction using random XY translations of 50–100 μm and 100–200 μm (Tables S2 and S3).

Additional file 2

File name: Supplementary_Video_S1.mp4

File format: MP4

Title: Time-lapse visualization of fungal network reconstruction.

Description: Shows the temporal evolution of the 25-hour fungal network reconstruction presented in Figure 10, including depth-coded reconstructions, representative bright-field microscopy planes, and OBB focus predictions.

References

1. Egbuta MA, Mwanza M, Babalola OO. Health Risks Associated with Exposure to Filamentous Fungi. Int J Environ Res Public Health. 2017 Jul 4;14(7):719. doi:10.3390/ijerph14070719 PubMed PMID: 28677641; PubMed Central PMCID: PMC5551157.

2. Cantrell SA, Dianese JC, Fell J, Gunde-Cimerman N, Zalar P. Unusual fungal niches. Mycologia. 2011 Nov 1;103(6):1161–74. doi:10.3852/11-108

3. Mycelial Interconnectedness | SpringerLink [Internet]. [cited 2025 Nov 22]. Available from: https://link.springer.com/chapter/10.1007/978-0-585-27576-5_2

4. Fricker MD, Heaton LLM, Jones NS, Boddy L. The Mycelium as a Network. Microbiol Spectr. 2017 May;5(3). doi:10.1128/microbiolspec.FUNK-0033-2017 PubMed PMID: 28524023; PubMed Central PMCID: PMC11687498.

5. Heaton L, Obara B, Grau V, Jones N, Nakagaki T, Boddy L, et al. Analysis of fungal networks. Fungal Biol Rev. 2012 Apr;26(1):12–29. doi:10.1016/j.fbr.2012.02.001

6. Lew RR. How does a hypha grow? The biophysics of pressurized growth in fungi. Nat Rev Microbiol. 2011 Jul;9(7):509–18. doi:10.1038/nrmicro2591

7. Boddy L, Wood J, Redman E, Hynes J, Fricker MD. Fungal network responses to grazing. Fungal Genet Biol. 2010 Jun;47(6):522–30. doi:10.1016/j.fgb.2010.01.006

8. Treseder KK, Lennon JT. Fungal traits that drive ecosystem dynamics on land. Microbiol Mol Biol Rev MMBR. 2015 Jun;79(2):243–62. doi:10.1128/MMBR.00001-15 PubMed PMID: 25971588; PubMed Central PMCID: PMC4429240.

9. Van Der Heijden MGA, Horton TR. Socialism in soil? The importance of mycorrhizal fungal networks for facilitation in natural ecosystems. J Ecol. 2009;97(6):1139–50. doi:10.1111/j.1365-2745.2009.01570.x

10. Magkourilou E, Bell CA, Daniell TJ, Field KJ. The functionality of arbuscular mycorrhizal networks across scales of experimental complexity and ecological relevance. Funct Ecol. 2025 Jun;39(6):1384–99. doi:10.1111/1365-2435.14618

11. Spatafora JW, Chang Y, Benny GL, Lazarus K, Smith ME, Berbee ML, et al. A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. Mycologia. 2016 Sep 1;108(5):1028–46. doi:10.3852/16-042 PubMed PMID: 27738200.

12. Bases of the obligate biotrophy of arbuscular mycorrhizal fungi | SpringerLink [Internet]. [cited 2025 Nov 23]. Available from: https://link.springer.com/chapter/10.1007/978-3-0348-8117-3_3

13. Smith SE, Read DJ. Mycorrhizal Symbiosis. Academic Press; 2010. 815 p.

14. Rhodes CJ. The Whispering World of Plants: 'The Wood Wide Web.' Sci Prog. 2017 Sep 1;100(3):331–7. doi:10.3184/003685017X14968299580423

15. Gorzelak MA, Asay AK, Pickles BJ, Simard SW. Inter-plant communication through mycorrhizal networks mediates complex adaptive behaviour in plant communities. AoB PLANTS. 2015 Jan 1;7:plv050. doi:10.1093/aobpla/plv050

16. Babikova Z, Gilbert L, Bruce TJA, Birkett M, Caulfield JC, Woodcock C, et al. Underground signals carried through common mycelial networks warn neighbouring plants of aphid attack. Ecol Lett. 2013;16(7):835–43. doi:10.1111/ele.12115

17. Symanczik S, Lehmann MF, Wiemken A, Boller T, Courty PE. Effects of two contrasted arbuscular mycorrhizal fungal isolates on nutrient uptake by Sorghum bicolor under drought. Mycorrhiza. 2018 Nov;28(8):779–85. doi:10.1007/s00572-018-0853-9

18. Ruiz-Lozano JM, Aroca R, Zamarreño ÁM, Molina S, Andreo-Jiménez B, Porcel R, et al. Arbuscular mycorrhizal symbiosis induces strigolactone biosynthesis under drought and improves drought tolerance in lettuce and tomato. Plant Cell Environ. 2016;39(2):441–52. doi:10.1111/pce.12631

19. Sikes BA, Cottenie K, Klironomos JN. Plant and fungal identity determines pathogen protection of plant roots by arbuscular mycorrhizas. J Ecol. 2009;97(6):1274–80. doi:10.1111/j.1365-2745.2009.01557.x

20. Koricheva J, Gange AC, Jones T. Effects of mycorrhizal fungi on insect herbivores: a meta-analysis. Ecology. 2009;90(8):2088–97. doi:10.1890/08-1555.1

21. Oyarte Galvez L, Bisot C, Bourrianne P, Cargill R, Klein M, Van Son M, et al. A travelling-wave strategy for plant–fungal trade. Nature. 2025 Mar 6;639(8053):172–80. doi:10.1038/s41586-025-08614-x

22. Aguilar-Trigueros CA, Boddy L, Rillig MC, Fricker MD. Network traits predict ecological strategies in fungi. ISME Commun. 2022 Dec 1;2(1):1. doi:10.1038/s43705-021-00085-1

1089 23. Martínez-Galicia E, Fernanda Flores Enríquez A, Puga A, Gutiérrez–Medina B. Analysis
1090 of the emerging physical network in young mycelia. Fungal Genet Biol. 2023
1091 Oct;168:103823. doi:10.1016/j.fgb.2023.103823

1092 24. Dikec J, Olivier A, Bobée C, D'Angelo Y, Catellier R, David P, et al. Hyphal network
1093 whole field imaging allows for accurate estimation of anastomosis rates and branching
1094 dynamics of the filamentous fungus Podospora anserina. Sci Rep. 2020 Feb 21;10(1):1.
1095 doi:10.1038/s41598-020-57808-y

1096 25. Schmideder S, Barthel L, Friedrich T, Thalhammer M, Kovačević T, Niessen L, et al.
1097 An X-ray microtomography-based method for detailed analysis of the three-dimensional
1098 morphology of fungal pellets. Biotechnol Bioeng. 2019 Jun;116(6):1355–65.
1099 doi:10.1002/bit.26956

1100 26. Müller H, Deffur C, Schmideder S, Barthel L, Friedrich T, Mirlach L, et al. Synchrotron
1101 radiation-based microcomputed tomography for three-dimensional growth analysis of
1102 Aspergillus niger pellets. Biotechnol Bioeng. 2023;120(11):3244–60.
1103 doi:10.1002/bit.28506

1104 27. Obara B, Grau V, Fricker MD. A bioimage informatics approach to automatically extract
1105 complex fungal networks. Bioinformatics. 2012 Sep 15;28(18):2374–81.
1106 doi:10.1093/bioinformatics/bts364

1107 28. Vidal-Diez De Ulzurrun G, Huang TY, Chang CW, Lin HC, Hsueh YP. Fungal feature
1108 tracker (FFT): A tool for quantitatively characterizing the morphology and growth of
1109 filamentous fungi. Poisot T, editor. PLOS Comput Biol. 2019 Oct 31;15(10):10.
1110 doi:10.1371/journal.pcbi.1007428

1111 29. Fricker MD, Akita D, Heaton LL, Jones N, Obara B, Nakagaki T. Automated analysis
1112 of Physarum network structure and dynamics. J Phys Appl Phys. 2017 Jun;50(25):254005.
1113 doi:10.1088/1361-6463/aa72b9

1114 30. Zhang Z, Nishimura Y, Kanchanawong P. Extracting microtubule networks from
1115 superresolution single-molecule localization microscopy data. Lippincott-Schwartz J,
1116 editor. Mol Biol Cell. 2017 Jan 15;28(2):333–45. doi:10.1091/mbc.e16-06-0421

1117 31. Zhang H, Söderholm N, Sandblad L, Wiklund K, Andersson M. DSeg: A Dynamic
1118 Image Segmentation Program to Extract Backbone Patterns for Filamentous Bacteria and
1119 Hyphae Structures. Microsc Microanal. 2019 Jun;25(3):711–9.
1120 doi:10.1017/S1431927619000308

1121 32. Yu F, Liu M, Chen W, Wen H, Wang Y, Zeng T. Automatic Repair of 3-D Neuron
1122 Reconstruction Based on Topological Feature Points and an MOST-Based Repairer. IEEE
1123 Trans Instrum Meas. 2021;70:1–13. doi:10.1109/TIM.2020.3033057

33. Mais L, Hirsch P, Managan C, Kandarpa R, Rumberger JL, Reinke A, et al. FISBe: A Real-World Benchmark Dataset for Instance Segmentation of Long-Range thin Filamentous Structures. In: 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) [Internet]. Seattle, WA, USA: IEEE; 2024 [cited 2024 Nov 11]. p. 22249–59. Available from: https://ieeexplore.ieee.org/document/10656899/ doi:10.1109/CVPR52733.2024.02100

34. Su H, Gao L, Wang Z, Yu Y, Hong J, Gao Y. A Hierarchical Full-Resolution Fusion Network and Topology-Aware Connectivity Booster for Retinal Vessel Segmentation. IEEE Trans Instrum Meas. 2024;73:1–16. doi:10.1109/TIM.2024.3411133

35. Luo X, Peng L, Ke Z, Lin J, Yu Z. PA-Net: A hybrid architecture for retinal vessel segmentation. Pattern Recognit. 2025 May;161:111254. doi:10.1016/j.patcog.2024.111254

36. Dirnberger M, Kehl T, Neumann A. NEFI: Network Extraction From Images. Sci Rep. 2015 Nov 2;5(1):15669. doi:10.1038/srep15669

37. Sten O, Del Dottore E, Ronzan M, Pugno N, Mazzolai B. Three-Dimensional Graph Reconstruction of Filamentous Structures from Z-Stack Images [Internet]. SSRN; 2024 [cited 2025 Dec 3]. Available from: https://www.ssrn.com/abstract=5039706 doi:10.2139/ssrn.5039706

38. Hardy N, Moreaud M, Guillaume D, Augier F, Nienow A, Béal C, et al. Advanced digital image analysis method dedicated to the characterization of the morphology of filamentous fungus. J Microsc. 2017 May;266(2):2. doi:10.1111/jmi.12523

39. De Ligne L, Vidal-Diez De Ulzurrun G, Baetens JM, Van Den Bulcke J, Van Acker J, De Baets B. Analysis of spatio-temporal fungal growth dynamics under different environmental conditions. IMA Fungus. 2019 Dec;10(1):7. doi:10.1186/s43008-019-0009-3

40. Ledoux C, Chapeland-Leclerc F, Ruprich-Robert G, Bobée C, Lalanne C, Herbert É, et al. Prediction and experimental evidence of different growth phases of the Podospora anserina hyphal network. Sci Rep. 2023 May 25;13(1):8501. doi:10.1038/s41598-023-35327-w

42. Brunk M, Sputh S, Doose S, Van De Linde S, Terpitz U. HyphaTracker: An ImageJ toolbox for time-resolved analysis of spore germination in filamentous fungi. Sci Rep. 2018 Jan 12;8(1):1. doi:10.1038/s41598-017-19103-1

43. Sánchez-Orellana G, Casas-Flores S, Gutiérrez–Medina B. Automated, continuous video microscopy tracking of hyphal growth. Fungal Genet Biol. 2019 Feb;123:25–32. doi:10.1016/j.fgb.2018.11.006

44. Sten O, Del Dottore E, Pugno N, Mazzolai B. A ridge-based detection algorithm with filament overlap identification for 2D mycelium network analysis. Ecol Inform. 2024 Sep;82:102670. doi:10.1016/j.ecoinf.2024.102670

45. Wang Y, Jia H, Jia P, Chen K, Zhang X. A novel algorithm for three-dimensional shape reconstruction for microscopic objects based on shape from focus. Opt Laser Technol. 2024 Jan;168:109931. doi:10.1016/j.optlastec.2023.109931

46. Pertuz S, Puig D, Garcia MA. Reliability measure for shape-from-focus. Image Vis Comput. 2013 Oct;31(10):725–34. doi:10.1016/j.imavis.2013.07.005

47. Pertuz S, Puig D, Garcia MA. Analysis of focus measure operators for shape-from-focus. Pattern Recognit. 2013 May;46(5):1415–32. doi:10.1016/j.patcog.2012.11.011

48. Østerlund I, Persson S, Nikoloski Z. Tracing and tracking filamentous structures across scales: A systematic review. Comput Struct Biotechnol J. 2023;21:452–62. doi:10.1016/j.csbj.2022.12.023

49. Shin SY, Lee S, Yun ID, Lee KM. Deep vessel segmentation by learning graphical connectivity. Med Image Anal. 2019 Dec;58:101556. doi:10.1016/j.media.2019.101556

50. Tahir W, Kura S, Zhu J, Cheng X, Damseh R, Tadesse F, et al. Anatomical modeling of brain vasculature in two-photon microscopy by generalizable deep learning. BME Front. 2020;2020:8620932. doi:10.34133/2020/8620932.

51. Liu Y, Kolagunda A, Treible W, Nedo A, Caplan J, Kambhamettu C. Intersection to Overpass: Instance Segmentation on Filamentous Structures With an Orientation-Aware Neural Network and Terminus Pairing Algorithm. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) [Internet]. Long Beach, CA, USA: IEEE; 2019 [cited 2026 Jan 17]. p. 125–33. Available from: https://ieeexplore.ieee.org/document/9025494/ doi:10.1109/CVPRW.2019.00021

52. Smith AG, Han E, Petersen J, Olsen NAF, Giese C, Athmann M, et al. RootPainter: deep learning segmentation of biological images with corrective annotation. New Phytol. 2022;236(2):774–91. doi:10.1111/nph.18387

53. Pound MP, Atkinson JA, Townsend AJ, Wilson MH, Griffiths M, Jackson AS, et al. Deep machine learning provides state-of-the-art performance in image-based plant phenotyping. GigaScience. 2017 Oct 1;6(10):gix083. doi:10.1093/gigascience/gix083

54. Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY, et al. SSD: Single Shot MultiBox Detector. In: Leibe B, Matas J, Sebe N, Welling M, editors. Computer Vision – ECCV 2016 [Internet]. Cham: Springer International Publishing; 2016 [cited 2026 Jan 18]. p. 21–37. (Lecture Notes in Computer Science). Available from: http://link.springer.com/10.1007/978-3-319-46448-0_2 doi:10.1007/978-3-319-46448-0_2

55. Zand M, Etemad A, Greenspan M. Oriented Bounding Boxes for Small and Freely Rotated Objects. IEEE Trans Geosci Remote Sens. 2022;60:1–15. doi:10.1109/TGRS.2021.3076050

56. Yi J, Wu P, Liu B, Huang Q, Qu H, Metaxas D. Oriented Object Detection in Aerial Images with Box Boundary-Aware Vectors. In: 2021 IEEE Winter Conference on Applications of Computer Vision (WACV) [Internet]. Waikoloa, HI, USA: IEEE; 2021 [cited 2026 Jan 19]. p. 2149–58. Available from: https://ieeexplore.ieee.org/document/9423055/ doi:10.1109/WACV48630.2021.00220

57. Vijayakumar A, Vairavasundaram S. YOLO-based Object Detection Models: A Review and its Applications. Multimed Tools Appl. 2024 Mar 14;83(35):83535–74. doi:10.1007/s11042-024-18872-y

58. Domínguez-Jiménez DY, RONZAN M, Del Dottore E, Mazzolai B. Bright-Field Microscopy Dataset of Fungal Filament Networks for Detection, Focus Modeling, and 3D Reconstruction [Internet]. 2026 Oct 4. doi:10.5281/zenodo.23135747

59. Ultralytics. Ultralytics YOLO [Internet]. GitHub. Available from: https://github.com/ultralytics/ultralytics/tree/opts-args. Accessed 4 Feb 2026.

60. Yang X, Yan J, Ming Q, Wang W, Zhang X, Tian Q. Rethinking Rotated Object Detection with Gaussian Wasserstein Distance Loss. In: Proceedings of the 38th International Conference on Machine Learning [Internet]. PMLR; 2021 [cited 2026 Feb 4]. p. 11830–41. Available from: https://proceedings.mlr.press/v139/yang21l.html

61. Staal J, Abramoff MD, Niemeijer M, Viergever MA, van Ginneken B. Ridge-based vessel segmentation in color images of the retina. IEEE Trans Med Imaging. 2004 Apr;23(4):501–9. doi:10.1109/TMI.2004.825627

62. Sasamura K, Ohki-Hamazaki H, Sugihara I. Morphology of the olivocerebellar projection of the chick: An axonal reconstruction study. J Comp Neurol. 2013;521(14):3321–39. doi:10.1002/cne.23352

63. North BV, Curtis D, Sham PC. A Note on the Calculation of Empirical P Values from Monte Carlo Procedures. Am J Hum Genet. 2002 Aug 1;71(2):439–41. doi:10.1086/341527 PubMed PMID: 12111669.

64. Xiao ZK, Yang GY, Yang X, Mu TJ, Yan J, Hu SM. Theoretically Achieving Continuous Representation of Oriented Bounding Boxes. In: 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) [Internet]. Seattle, WA, USA: IEEE; 2024 [cited 2026 Mar 30]. p. 16912–22. Available from: https://ieeexplore.ieee.org/document/10657424/ doi:10.1109/CVPR52733.2024.01600

65. Liu ZL, Wang HW, Weng LW, Yang YY. HRSC2016 [Internet]. IEEE DataPort; [cited 2026 Mar 30]. Available from: https://dx.doi.org/10.21227/rgx1-sh71 doi:10.21227/RGX1-SH71

66. Sun X, Wang P, Yan Z, Xu F, Wang R, Diao W, et al. FAIR1M: A benchmark dataset for fine-grained object recognition in high-resolution remote sensing imagery. ISPRS J Photogramm Remote Sens. 2022 Feb;184:116–30. doi:10.1016/j.isprsjprs.2021.12.004

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial.docx
- SupplementaryVideoS1.mp4