



Development of a method for the detection of peat in growing media

Bern, 27.01.2025

Commissioned by the Federal Office for the Environment (FOEN)

Imprint

Commissioned by:

Federal Office for the Environment (FOEN), Economics and Innovation Division, CH-3003 Bern

The FOEN is an agency of the Federal Department of the Environment, Transport, Energy and Communications (DETEC).

Contractor:

Federal Institute for Metrology METAS

Authors:

Lena Märki¹, Serge Zaugg¹, Camille Vögeli², Edward Mitchell²

¹ Federal Institute for Metrology METAS, Lindenweg 50, 3003 Bern-Wabern, Switzerland

² University of Neuchâtel, Laboratory of Soil Biodiversity, Institute of Biology, Rue Emile Argand 11, 2000 Neuchâtel, Switzerland

FOEN support:

Laura Tschümperlin, Josef Känzig

Note:

This study/report was prepared under contract to the Federal Office for the Environment (FOEN). The contractor bears sole responsibility for the content.

Supplementary material:

File name	Description
grid_search_Archerella_sp.png	Ranked <i>Archerella</i> candidates from processing 1600 grid search images in the test set.
grid_and_active_Archerella_sp.png	Ranked <i>Archerella</i> candidates from processing 1872 grid + active search images in the test set.
grid_and_active_Amphitrema_sp.png	Ranked <i>Amphitrema</i> candidates from processing 1872 grid + active search images in the test set.
grid_and_active_Assulina_sp.png	Ranked <i>Assulina</i> candidates from processing 1872 grid + active search images in the test set.

Definitions and acronyms

Substrate Growing media	We use substrate and growing media as synonyms. The term substrate is also used for commercial products composed of pure peat. It is used for plant growth in the professional sector (especially for seedlings and potted plants) and can be bought in bags for gardening.
Commercial peat	We use the term commercial peat for describing peat that has been extracted and processed for the production of growing media.
Testate amoebae	Unicellular microorganisms. Some species of testate amoebae are specialized to the specific and extreme conditions (low pH, low nutrients) of peatlands and uniquely live in peat soils.
<i>Sphagnum</i>	Genus of mosses that occur preferentially in peatlands. They have the capacity to store water in their living or dead cells, to lower pH and are highly efficient in retaining nutrients.
YOLO v8	YOLO is an open-source deep learning model for object detection in images. We use the version 8 (v8) of YOLO.
EMDS7 dataset	A public research dataset with microscopy images from Environmental microorganism [60].
Crops	An automatically extracted rectangular region of an image.
Object detection	A process that identifies the position and class of object inside pictures.
Model	In this document this refers to a complex Deep Neural Network of the YOLO family that can be trained with data to perform object detection.
Prediction module	A process implemented in the Python programming environment that ingest multiple microscopy images and returns a list of small images that should contain the target amoebae.
Score	A value between 0 and 1 that is automatically predicted by an algorithm an tells how sure the algorithm is of its prediction.
Ensemble score	A derived score obtained by combing the score of several models. In our case the mean was used to compute the ensemble score.
Precision	Precision is the fraction of relevant instances among all the retrieved instances.
Recall	Recall is the fraction of relevant instances that were retrieved from all relevant instances.
Final training set	5121 images from 33 samples that were used to train all models used in the Prediction module
Test Set	2415 images from 16 samples the are independent of Final training set
Average Precision	See AP
AP	Average Precision – The standard metric to measure detection performance of object detection algorithms.
DNN	Deep Neural Network
DA	Data Augmentation
IoU	Intersection over Union
ML	Machine Learning
OD	Object Detection

Table of Content

1	Summaries	7
1.1	English	7
1.2	Deutsch	8
1.3	Français	9
2	Introduction	10
2.1	Background	10
2.2	Aim of this project	10
2.3	Secondary method: long-chain n-alkanes	11
2.4	Testate amoebae	12
2.5	Related deep learning work	13
3	Method	13
3.1	Selection of peat-indicator testate amoebae	13
3.2	Image acquisition	15
3.2.1	Sample preparation	15
3.2.2	Grid search	15
3.2.3	Active search	15
3.3	Image preparation	16
3.4	Data	16
3.4.1	Data sets	16
3.4.2	Data splitting	16
3.5	Model training	18
3.5.1	Preparation of training data	18
3.5.2	Models and training sessions	18
3.6	Performance evaluation	18
3.7	Final prediction module via ensemble	18
4	Results	19
4.1	Testate amoebae in commercial peat samples	19
4.2	Testate amoebae in commercial non-peat samples	20
4.3	Model comparison from cross validation	22
4.4	Performance with test data	23
4.5	Illustration of module with test data	23
4.5.1	<i>Archerella</i> grid only	24
4.5.2	<i>Archerella</i> grid and active	24
4.5.3	<i>Amphitrema</i> grid and active	24
4.5.4	<i>Assulina</i> grid and active	24
5	Discussion	28
5.1	Shells preserved in commercial peat products	28
5.2	Commercial peat-free substrates and components	28
5.3	Strengths and limitations	28
6	Outlook	29

6.1	Large-scale analysis with an automated microscope	29
6.2	Need for automated microscope.....	29
6.3	Development of a method to quantify peat percentage.....	29
6.4	Cultivated <i>Sphagnum</i> moss.....	29
6.5	Geographical scope of the method	30
6.6	Extension to other applications.....	30
7	Acknowledgments.....	30
8	Conclusions	31
9	References	32
10	Appendix	37

1 Summaries

1.1 English

Peat is formed by the slow accumulation of organic material in water-saturated soils. Peatlands play a key role in the global carbon cycle, biodiversity conservation and hydrological regulation. The degradation of peatlands through drainage and peat extraction contributes to greenhouse gas emissions, biodiversity loss and affects water resources. Most extracted peat is used for production of energy or as substrate for plants (growing media). Switzerland, as well as other countries, aims to reduce the use of peat in substrates. This requires the development of tools that measure the peat content of substrate.

We propose a method for peat detection based on the identification of peat-specific microorganisms called testate amoebae. This group of unicellular microorganisms produce shells. These shells can be preserved in peat for thousands to millions of years. The taxa (species or group of species) of testate amoebae can be identified by experts based on microscopy images thanks to the morphological characteristics of their shells. Some species of testate amoebae are specialized for the conditions in peat soils and uniquely occur in peatlands. These peat-specific species are used here as bioindicators for the detection of peat in substrates.

Since identifying testate amoebae in microscopy images is tedious and few experts exist, we propose a decision support system based on machine learning (ML). Our system detects testate amoebae in large collections of microscopy images. A specific dataset of images from commercial substrates was acquired and annotated for the ML-system. A dedicated data augmentation procedure was developed to mitigate the small training sample size. YOLOv8 models for object detection were trained and then evaluated with an independent test set from commercial substrates.

We found that the shells of two peat-specific taxa (*Archerella* sp and *Amphitrema* sp) were well preserved in commercial substrates and can serve as indicators of peat presence. With the test set, detection performance given as average precision* was good with values above 0.8 for *Archerella* sp. and above 0.7 for *Amphitrema* sp. Our decision support system can process thousands of images in a few minutes, helping a human operator to quickly decide if peat is present in a sample. This method has the potential to be further developed for an estimation of the proportion of peat in a substrate.

*The detection performance of an ML algorithm is typically measured using the average precision. A value close to zero indicates poor performance. Values between 0.5 and 1.0 indicate useful to very good performance.

1.2 Deutsch

Torf entsteht durch die langsame Ansammlung von organischem Material in wassergesättigten Böden. Torfgebiete spielen eine Schlüsselrolle für den globalen Kohlenstoffkreislauf, die Erhaltung der Biodiversität und die Regulierung des Wasserhaushalts. Die Degradierung von Torfgebieten durch Entwässerung und Torfabbau trägt zu Treibhausgasemissionen und zum Verlust der biologischen Vielfalt bei und beeinträchtigt die Pufferrolle in der hydrologischen Regulierung. Torf wird am häufigsten zur Energiegewinnung oder zur Verwendung als Substrat für Pflanzen abgebaut. Die Schweiz und andere Länder wollen die Verwendung von Torf für Substrate reduzieren. Dies erfordert die Entwicklung von Methoden zur Messung des Torfgehalts in Substraten.

Wir schlagen eine Methode zum Torfnachweis vor, die auf der Identifizierung von torfspezifischen Mikroorganismen beruht, den so genannten Schalenamöben. Diese Gruppe von einzelligen Mikroorganismen produziert Schalen. Diese Schalen können im Torf über Tausende bis Millionen von Jahren erhalten bleiben. Die Taxa (Arten oder Artengruppen) der Schalenamöben können von Fachleuten in Mikroskopie Bildern anhand der morphologischen Merkmale bestimmt werden. Einige Arten von Schalenamöben sind auf die Bedingungen von Torfböden spezialisiert und kommen nur in Torfgebieten vor. Diese torfspezifischen Arten werden hier als Bioindikatoren für den Nachweis von Torf in Substraten verwendet.

Da die Identifizierung von Schalenamöben in Mikroskopie Bildern mühsam ist und es nur wenige Experten gibt, schlagen wir ein System zur Entscheidungshilfe vor, das auf maschinellem Lernen (ML) basiert. Unser System erkennt Schalenamöben in grossen Sammlungen von Mikroskopie Bildern. Ein spezieller Datensatz mit Mikroskopie-Bildern von kommerziellen Substraten wurde für die Entwicklung des ML-Systems erfasst und annotiert. Ein Verfahren zur Bildtransformation wurde entwickelt, um die geringe Grösse der Trainingsstichprobe auszugleichen. YOLOv8 Modelle für die Objekterkennung wurden trainiert und anschliessend mit einem unabhängigen Testdatensatz aus kommerziellen Substraten evaluiert.

Unsere Untersuchung ergab, dass die Schalen von zwei torfspezifischen Taxa (*Archerella* sp. und *Amphitrema* sp.) in kommerziellen Substraten gut erhalten sind und als Indikatoren für das Vorhandensein von Torf dienen können. Mit dem Testdatensatz war die Durchschnittliche Präzision* (Erkennungsleistung) gut mit Werten über 0.8 für *Archerella* sp. und über 0.7 für *Amphitrema* sp. Unser System zur Entscheidungshilfe kann Tausende von Bildern in wenigen Minuten verarbeiten und menschlichen Mitarbeiterinnen und Mitarbeiter helfen, schnell zu entscheiden, ob in einer Probe Torf vorhanden ist. Diese Methode könnte weiterentwickelt werden, um den Torfanteil in einem Substrat zu schätzen.

*Die Erkennungsleistung eines ML-Algorithmus misst man standardmässig mit der Durchschnittliche Präzision (aus dem English average precision). Ein Wert nahe Null weist auf eine schlechte Leistung hin. Werte zwischen 0.5 und 1.0 zeigen eine nützliche bis sehr gute Leistung an.

1.3 Français

La tourbe est formée par l'accumulation lente de matière organique dans des sols saturés en eau. Les tourbières jouent un rôle clé dans le cycle du carbone, la conservation de la biodiversité et la régulation hydrologique. La dégradation des tourbières par le drainage et l'extraction de tourbe contribue aux émissions de gaz à effet de serre, à la perte de biodiversité et affecte les ressources en eau. La majorité de la tourbe extraite est utilisée pour la production d'énergie ou comme substrat pour les plantes (milieux de culture). La Suisse, comme d'autres pays, vise à réduire l'utilisation de la tourbe comme substrat. Pour ce faire, il est nécessaire de développer des outils permettant de mesurer la teneur en tourbe des substrats.

Nous proposons une méthode de détection de la tourbe basée sur l'identification de microorganismes spécifiques à la tourbe appelés amibes à thèque. Ce groupe de micro-organismes unicellulaires produit des coquilles. Ces coquilles peuvent être préservées dans la tourbe pendant des milliers, voire des millions d'années. Les taxons (espèces ou groupes d'espèces) d'amibes à thèque peuvent être identifiés par des experts sur la base d'images de microscopie grâce aux caractéristiques morphologiques de leurs coquilles. Certaines espèces d'amibes sont spécialisées pour les conditions des sols tourbeux et ne se trouvent que dans les tourbières. Ces espèces spécifiques à la tourbe sont utilisées ici comme bioindicateurs pour la détection de la tourbe dans les substrats.

Comme l'identification de ces amibes dans les images de microscopie est fastidieuse et qu'il n'existe que peu d'experts, nous proposons un système d'aide à la décision basé sur l'apprentissage automatique. Notre système détecte les amibes à thèque dans de grandes collections d'images. Un jeu de données spécifique provenant de substrats commerciaux a été acquis et annoté. Une procédure d'augmentation des données a été développée pour permettre de travailler avec la petite taille de l'échantillon d'entraînement. Les modèles YOLOv8 de détection d'objets ont été entraînés puis évalués à l'aide d'un jeu de données indépendants provenant de substrats commerciaux.

Nous avons constaté que les coquilles de deux taxons spécifiques à la tourbe (*Archerella* sp et *Amphitrema* sp) étaient bien conservées dans les substrats commerciaux et pouvaient servir d'indicateurs de la présence de tourbe. La performance de détection, exprimée en précision moyenne*, était bonne avec des valeurs supérieures à 0.8 pour *Archerella* sp et supérieures à 0.7 pour *Amphitrema* sp. Notre système d'aide à la décision peut traiter des milliers d'images en quelques minutes, aidant ainsi un opérateur humain à décider rapidement si de la tourbe est présente dans un échantillon. Cette méthode pourrait être perfectionnée pour estimer la proportion de tourbe dans un substrat.

*La performance de reconnaissance d'un algorithme ML est généralement mesurée par la précision moyenne (en anglais average precision). Une valeur proche de zéro indique une mauvaise performance. Des valeurs entre 0.5 et 1.0 indiquent une performance utile à très bonne.

2 Introduction

2.1 Background

Peat is formed by the accumulation of more or less decomposed organic matter in mires (peat-rich fens, raised bogs, or tropical swamps such as forested or papyrus swamps) under hydromorphic and anaerobic conditions. Peatlands play a key role in the global carbon cycle as they act as important carbon sinks. As such, peatlands store around 30% of the soil organic carbon worldwide, although they only cover around 3% of the Earth's surface [1, 2]. However, this natural system of carbon storage is endangered by human activities. The degradation of peatlands through drainage and peat extraction leads to aerobic mineralization of the stored organic matter and thus to significant amounts of greenhouse gas emissions. It has been estimated that the destruction of peatlands accounts for about 5% of the global anthropogenic greenhouse gas emissions which corresponds to around 2 Gigatons of CO₂ per year [3, 4]. The conservation and rewetting of peatlands (and thus the restoration of their carbon storage capacity) are therefore among the most efficient actions to mitigate climate change [5]. Peatlands also form ecosystems of major importance for biodiversity conservation, which support numerous rare species adapted to the specific site conditions. Additionally, peatlands play key roles in hydrological regulation; by storing water during rainy periods, they act as buffers against floods and help maintaining baseline water flow during dry periods. As such, peatlands and other wetlands also contribute to stabilizing regional climate by reducing temperature extremes [6].

Thanks to the Rothenturm Initiative, peatlands are protected in Switzerland since 1987, and thus peat extraction is prohibited in the country. However, an estimated 500'000 m³ of peat is imported into Switzerland every year [68]. This peat is used mostly in substrates for professional vegetable, fruit and ornamental plant cultivation. Several natural characteristics of peat make it suitable for its use as an ingredient of substrate or as soil enhancer: its important capacity to store water, its structural stability due to the poorly decomposed *Sphagnum* moss debris, its low pH as well as its low levels of nutrients and pollutants [74]. With increasing awareness that peat resources are finite and ecologically valuable, several substrate producers aim to reduce their use of peat.

In order to reduce the import of peat, the Swiss Federal Council adopted the *Peat exit plan* in 2012. In a first phase, peat imports are to be reduced through voluntary measures. In a second phase, trade policy measures should be examined if necessary. By raising awareness among the public and declarations of intent (Memorandum of understanding) from substrate manufacturers and retailers, the peat content in substrates for hobby gardening has been reduced by more than 50% since 2015 [69].

The declaration and labelling of peat-free substrates is nowadays based on the traceability of supply chains. However, it is currently impossible to scientifically certify the absence of peat in a substrate. This is also true for the data collection of the peat import into Switzerland, which is until now based on questionnaires. The aim of this project was to fill this gap by developing and testing methods to detect peat in substrates.

2.2 Aim of this project

The objective of this study was to develop a method to detect peat in growing media. The primary focus was on the peat detection by the identification of peat-specific testate amoebae, as a preliminary study had shown the potential of this approach. The automation of testate amoeba identification should allow to batch-process images from multiple samples and extract small images (crops) of likely testate amoebae. This will be part of a decision support system where highly digested summaries of the crops will be available to human experts who take final decisions on peat presence with minimal time and effort. The method is expected to enable large scale monitoring of commercial substrates to monitor the presence of peat.

2.3 Secondary method: long-chain n-alkanes

A secondary peat-detection method based on long-chain n-alkanes was tested. Long-chain n-alkanes occur in the epicuticular wax layer of plants and are chemically inert. They remain in peat for a long time period without being decomposed. The composition of the n-alkane chain length varies depending on the plant species and several studies have suggested that Sphagnum mosses produce n-alkanes with characteristic chain lengths [70-72]. The aim was to use this characteristic n-alkanes of Sphagnum for the detection of peat. Long-chain alkanes can be determined by GC-MS (gas chromatography – mass spectrometry) analysis. As part of this project, the sample preparation as well as the GC-MS method were optimized. Subsequently, 48 peat-containing and peat-free samples were measured using GC-MS analysis. The data was used to classify the samples according to their peat presence both by calculating an alkane score and by applying a machine learning system.

From the chromatograms, we extracted 23 values as the areas of the peaks detected for the alkanes C11 to C33. A principal component analysis of these 23 values suggests that alkanes would be suitable for peat detection, as peat-containing and peat-free samples were mapped on different regions of the first two components (**Fig. 1**). By calculating the p-values of the individual alkanes, it was possible to show that the n-alkane chain lengths C31 and C33 in particular occur in statistically higher proportions in samples containing peat. A sum of the peak area ratios of the signals for C31 and C33 was therefore calculated to classify the samples. Using a decision threshold of 0.5 on this score (**Fig. 2**), 13 out of 17 pure peat samples and 2 out of 9 mixed samples could be distinguished from peat-free products (**Fig. 2**). As an outlook, a machine learning system was created by calculating a linear discriminant analysis, which has not yet been evaluated on a test set due to the small size of the data set (**Fig. 3**). It shows promising results, but they would have to be confirmed with an independent test set.

To conclude, we showed that peat detection through n-alkanes may be feasible. But to demonstrate it, more resources should be invested to gather and analyse a much larger dataset representative of all potential peat-substitution products. However, the n-alkanes method requires time-consuming sample preparation and the availability of expensive instruments (GC-MS). The microscopy method is based on highly specific testate amoeba biomarkers which allows to develop a predictive model without need of a large dataset with all potential peat-substitution products. For the above reason, the further development of the n-alkanes method was therefore not prioritized.

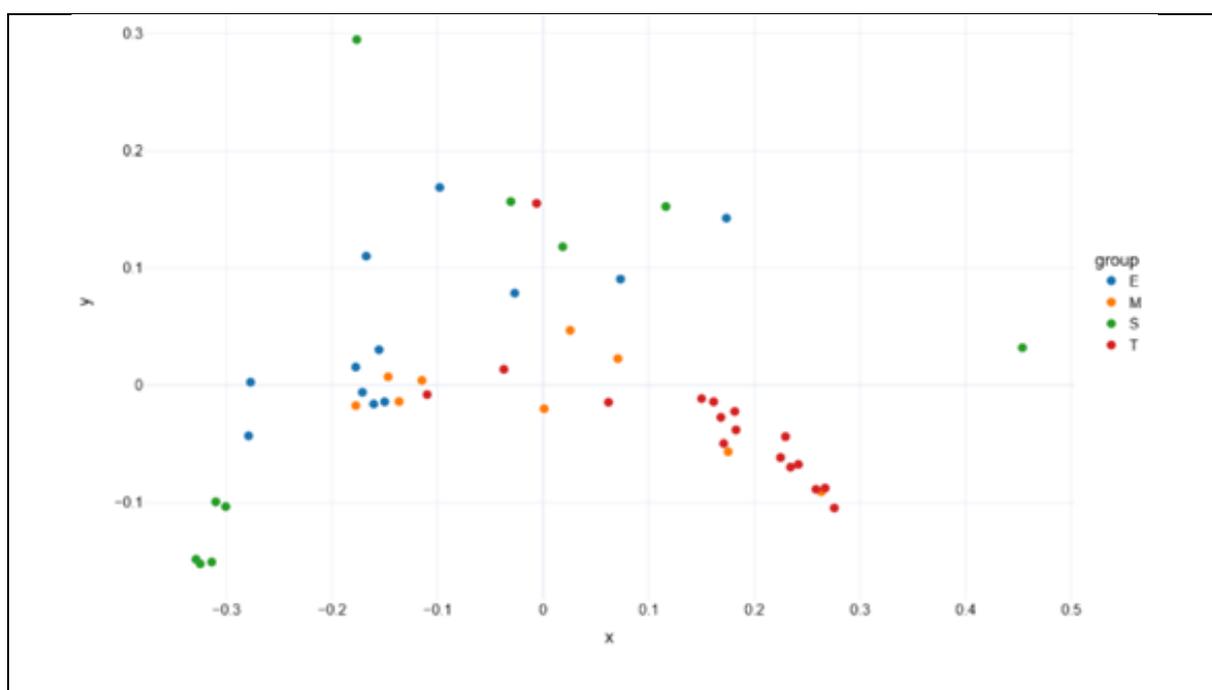


Figure 1. Two-dimensional PCA (76 % of the total variance) of the four sample groups: peat-free soil samples (E, blue), peat-free substitution samples (S, green), mixed samples containing peat (M, orange), pure peat (T, red).

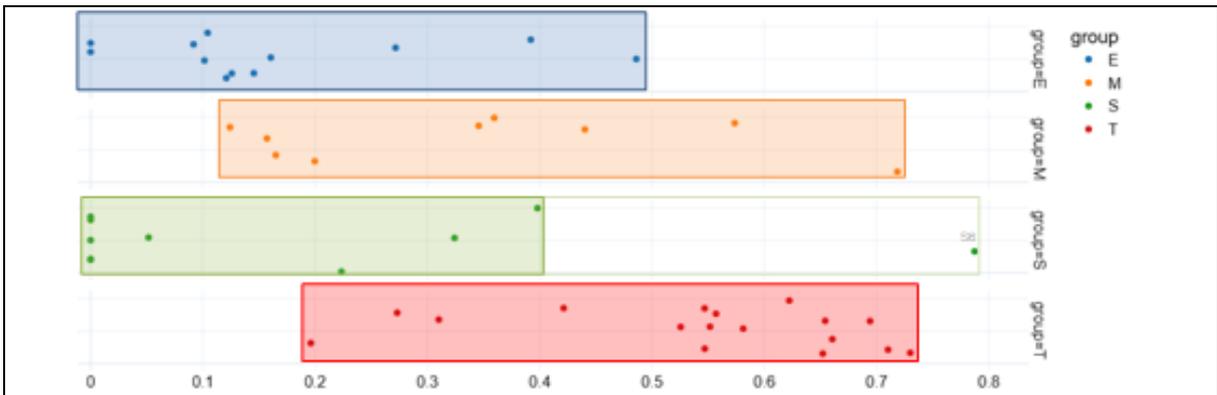


Figure 2. Illustration of the alkane score (C31+C33) from the data points. Soil samples (blue), mixed samples (orange), substitution samples (green), peat samples (red).

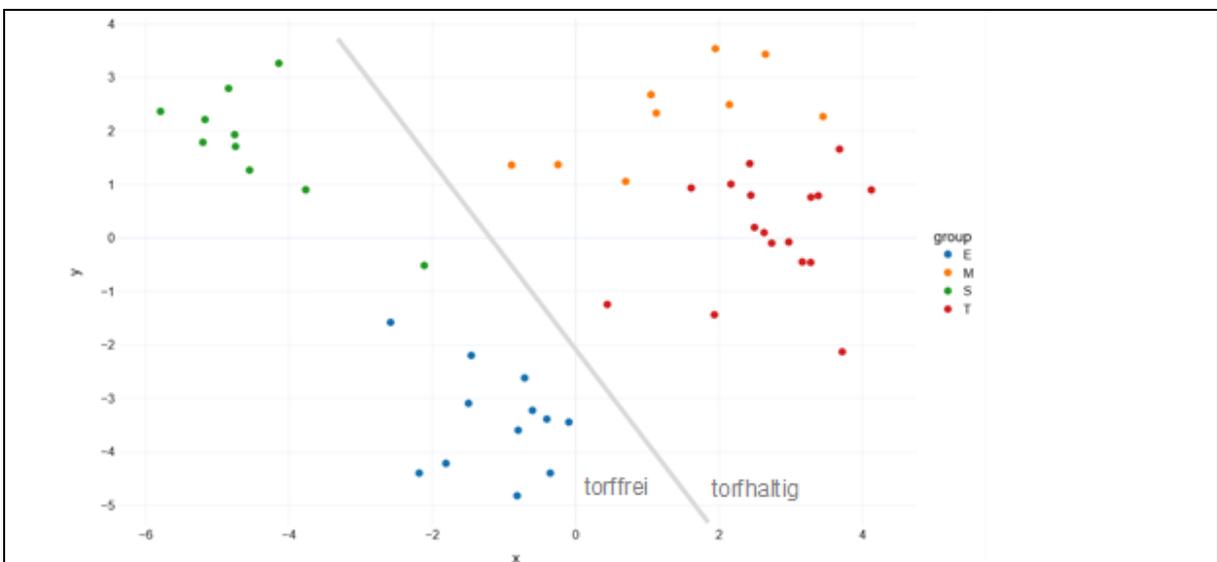


Figure 3. LDA of alkane proportions from 48 samples, including 12 soil samples (blue), 10 substitution samples (green), 9 mixed samples (orange) and 17 pure peat samples (red). A possible boundary (gray line) was manually drawn between the peat-containing and peat-free samples.

2.4 Testate amoebae

Testate amoebae are a common and diverse group of free-living amoeboid protists. The shell (called test) is either secreted (SiO₂, calcite, or protein) or built from recycled organic or mineral particles glued together with an organic cement and allows identification to species level [8, 9]. The shells remain after the death of the amoeba and under some conditions (anoxia, volcanic deposition) may preserve for millennia [10] to millions of years [11, 12] and even hundreds of millions of years [13, 14, 15]. Testate amoebae are commonly used as bioindicators of present and past environmental conditions, especially in peatlands where they are mostly used as hydrological indicators (water table depth) but also for pH and nutrient status [16, 17, 18], freshwater habitats [19, 20, 21, 22] and estuaries [23]. Testate amoebae are also used as bioindicators in lakes where they respond to nutrients, and heavy metal pollution [16, 22]. Testate amoebae are increasingly used to monitor peatland functioning [24, 25] and restoration success [26, 27, 28, 29, 30, 31, 32] as well as to assess the impact of forest management [33]. Peatlands are home to a high diversity of testate amoebae [34]. In a recent monograph, Bankov and Todorov [35] listed 175 testate amoeba species living in *Sphagnum* in Bulgaria. There are no compilations for testate amoeba diversity across

broader regions or globally at high taxonomic resolution, but it is very likely that the total diversity of testate amoebae existing in peatlands worldwide is well over 200 species. However, not all species listed as occurring in *Sphagnum* or in peatlands are restricted to these habitats. Many species found in peatlands may also be found in acidic forest litter or freshwater habitats. This may in part be due to the existence of several morphologically similar species within a given morphotype. A detailed analysis of such a species complex (*Nebela tincta* group) in the Jura Mountains revealed that closely related species differed in their ecology, some being specific to forested peatlands while others occurred preferentially in wetter and more nutrient rich habitats [36]. Still several taxa are clearly specific to *Sphagnum*-dominated peatlands, being frequent in *Sphagnum* and rare or absent from other habitats. This includes several mixotrophic taxa (i.e. the genera *Archerella* and *Amphitrema*, and the species *Heleopera sphagni* and *Placocista spinosa*) that harbour endosymbiotic green algae (*Chlorella*) [62]. This metabolism allows them to thrive in the nutrient-depleted habitats of peatlands.

2.5 Related deep learning work

Deep learning algorithms for image processing have seen a steep development in the past few years. Many mature algorithms are now available and have proven to achieve low error rates on difficult tasks. Deep learning algorithms have been successfully applied to microscopy images in a variety of health-related domains such as histopathology [37, 38, 39, 40], bacterial cultures [41] and blood parasites [42, 43, 44, 45]. Many deep learning methods have been developed for processing microscopy images in general (See [46, 47, 48] for reviews). A few studies have applied deep learning specifically to detect environmental microorganisms in microscopy images [49, 50, 51, 52]. We found only one study that focused on testate amoebae, but this was on activated sludge and not peatlands [53]. In many applied scenarios object detection (OD) models have proven their usefulness. A trained OD model can automatically predict rectangular regions of an image that contain the target together with a confidence value. The YOLO family of models for OD was introduced in 2016 [54] and has undergone a steep evolution since then ([55] for a review). YOLO models were originally developed to detect usual objects on photos (humans, dogs, cars, apples). They have proven to be very general and have been used in diverse scenarios, including microscopy. A detailed description of the model architecture can be found in [56]. For this study we used YOLOv8, which is well integrated in the Python ecosystem. Models from the YOLO family have been applied to microscopy images in various domains such as: microbes in industrial sludge [53], bacterial solutions in micro-fluidic chips [57], malaria parasites in blood [42, 43, 44, 45], small algae and diatoms [58, 59].

3 Method

3.1 Selection of peat-indicator testate amoebae

For the detection of peat, we are interested in species that occur exclusively in peatlands. We selected species characteristic for peat based on extensive data sets of testate amoebae community from Holarctic peatlands [63, 64]. We identified a set of species that together are found in most samples. Species were combined into classes (species group) that could be unambiguously identified by several experts. These classes were then used to annotate the data and to train the deep learning models. Taking into consideration the ease of identification, frequency of occurrence, and specificity of habitat, we selected 10 classes (**Table 1**). We also included species that are commonly found in peatlands but that are not peat indicators for an assessment of the method's more general ability to detect testate amoebae based on their morphology.

Archerella flavum was by far the most frequent taxon found in commercial peat samples. This genus also has a very typical shape, color, and appearance (**Fig. 5** and **6A**, and **Fig. A1** and **A2** in appendix). Due to its overall high frequency, degraded, folded or heavily masked *Archerella* shells were also observed. Thus, we defined an additional class for these (class *Archerella* degraded (**Fig. 6B**, **Fig. A3** in appendix). This allowed us to train and validate

models with the subset of clear examples (class *Archerella* sp). The trained model will therefore predict clear *Archerella* which will be easier to understand for end users in the proposed decision support process. For the other eight classes this distinction was not done due to their lower frequency in commercial samples.

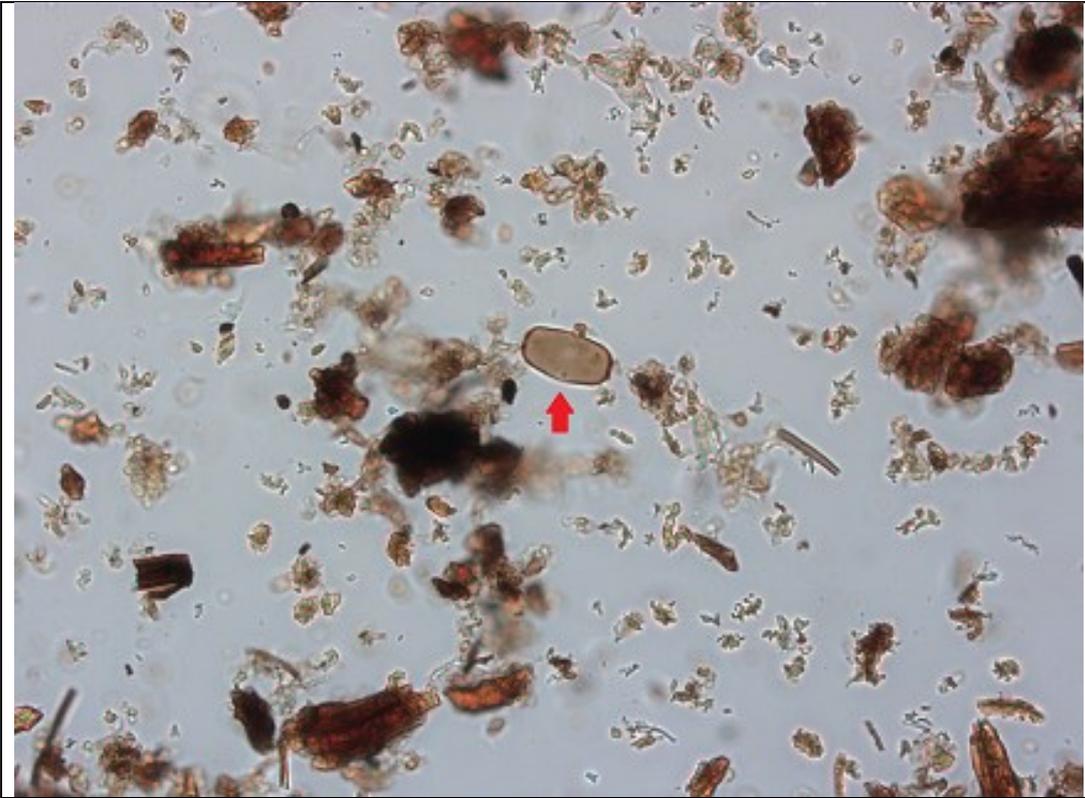


Figure 5. One image obtained with 20x magnification from a commercial peat sample. A shell of *Archerella* sp. is shown with a red arrow. Many plant residues are present all over the image. However, none of the plant debris can be recognised as belonging to *Sphagnum* mosses.

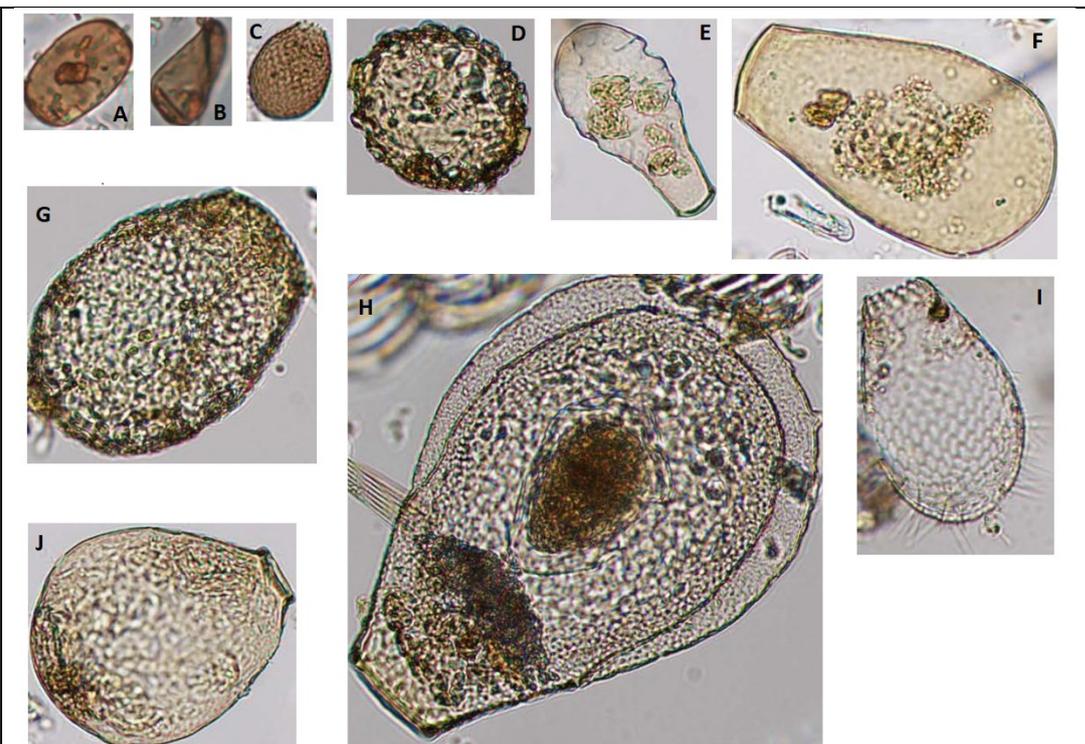


Figure 6. Individuals from the 10 testate amoeba species groups used in this study. (A) *Archerella* sp (B) *Archerella* degraded (C) *Assulina* sp (D) *Amphitrema* sp (E)

3.2 Image acquisition

3.2.1 Sample preparation

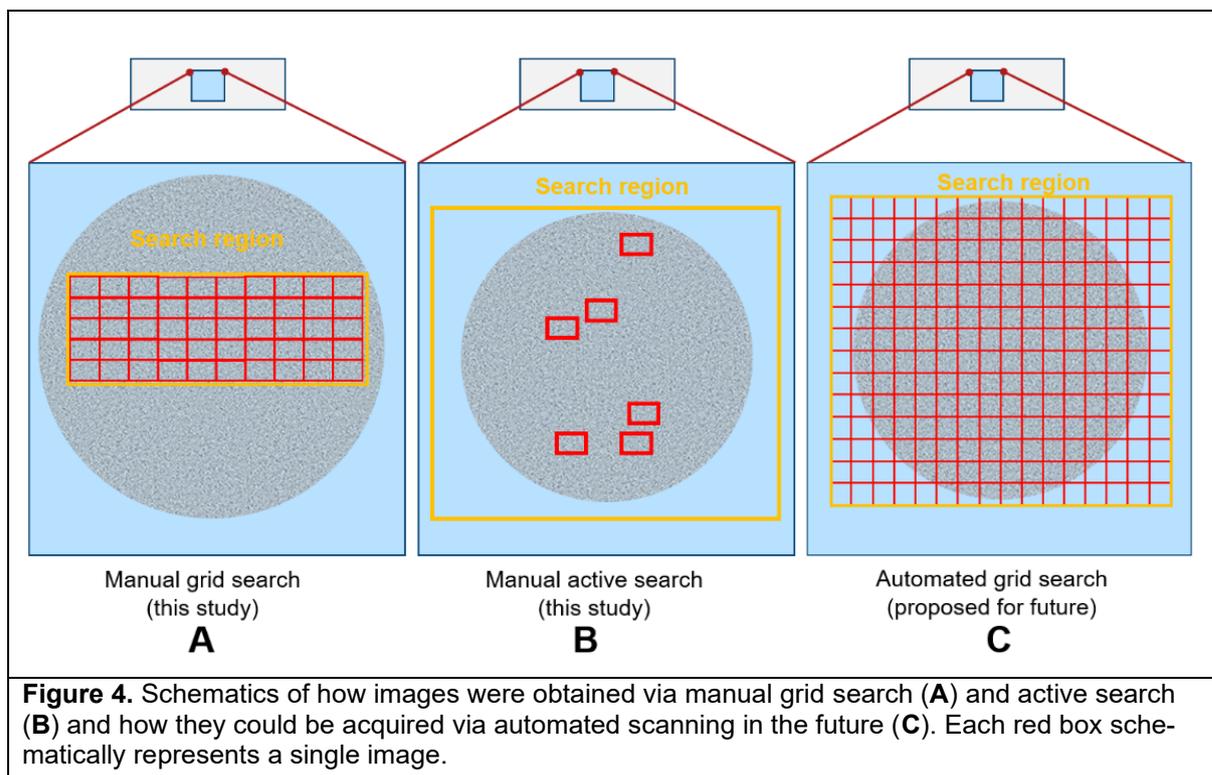
A small volume (ca. 5-10 cm³) of substrate was mixed with water, shaken for 1 minute in a wide screw-capped jar and filtered through a tea strainer. The material was then passed through an 80 µm mesh, which removes coarse particles with only marginal loss of testate amoebae, except for the largest species such as *Planocarina carinata*. The filtrate was left to settle overnight after which the clear supernatant was carefully poured off. The concentrate was then transferred to a tube. One drop of this concentrate was placed on a slide with a pipette and mixed with one drop of glycerol. Images were acquired under brightfield microscopy at 20x magnification with a camera mounted on the microscope and stored in TIFF format. For this project, as an automated microscope was not available, images were manually acquired. An early exploration showed that commercial peat samples contained a low density of testate amoeba shells. Thus, we defined two complementary image acquisition procedures: grid search and active search.

3.2.2 Grid search

A 5 by 10 grid of adjacent images are manually taken (**Fig. 4 A**). For each sample, 2 or 3 slides were imaged resulting in 100 or 150 images per sample. Most images typically do not contain any testate amoeba (only plant remains) and, when present, testate amoebae are generally not centred. The grid search mimics a realistic application scenario where data is acquired by an automatic scanning microscope (**Fig. 4 C**).

3.2.3 Active search

The whole slide is visually explored, but pictures are only taken when target amoebae are found and the amoebae are centred in the picture (**Fig. 4 B**). Therefore, all images contain at least one individual. This procedure was designed to capture all target amoebae present under a slide. In addition, additional images were taken with each observed amoeba placed either close to the bottom-left or top-right corner of the image.



3.3 Image preparation

A standard image width of 1728 pixels was defined. The target magnification was 20x, which corresponds to approximately 0.32 $\mu\text{m}/\text{pixel}$. All validation and test images were acquired with a 20x magnification by-design but some training images that had been acquired prior to project start had 40x magnification. The latter images were downsized to obtain a resolution of 0.32 $\mu\text{m}/\text{pixel}$ and they were then padded to reach the standard width of 1728 pixels. For padding we used images from the EMDS7 dataset [60]. The original EMDS7 images were greenish, so we converted them to black and white and then to a variety of soft colors to make them more diverse. All images were scrutinized by testate amoeba experts with a dedicated tool (Roboflow) and the position of the testate amoebae in the images was annotated in the form of bounding boxes. No distinction was made between dead or alive individuals because our focus was on detecting the shells. All amoeba that could be recognized were annotated even if they were degraded or masked. The annotated images were exported as JPG files in the YOLO format. A summary of classes is shown in **Table 1**.

3.4 Data

3.4.1 Data sets

Training: The training set was derived from 16 *Sphagnum* moss samples collected in peatlands (i.e., not commercially processed). The rationale for this choice is that natural samples contain the highest possible diversity and abundance of testate amoebae. These data were available prior to the study as part of the image repository of the Laboratory of Soil Biodiversity. These images were rich as they contained testate amoeba of many species (**Table 3, Table A1**). Often the amoebae were alive and presented in diverse natural colors. Note that these images are not representative of commercially processed substrate and were therefore used exclusively for training and not for validation or test.

Validation and test: The validation and test sets were derived exclusively from commercial substrates. They were sequentially acquired during this study (first validation set, second test set). In these images, the shells were often empty, more or less degraded and the coloration was less vivid than living specimen.

Validation: The validation set was derived from 17 samples taken from commercial substrates (10 peat, 1 *Sphagnum* moss, and 6 non-peat, **Table 2, Table A3**). These data are representative of commercially processed substrate and were either added to the training set or used for validation in a cross-validation procedure. From each sample, we prepared one suspension that was used to make 5 slides (2 for active search, 3 for grid search).

Test: The test set was derived from 16 samples taken from commercial substrates (12 peat, 4 non-peat, **Table 2, Table A4**). From each sample we prepared one suspension that was used to make 2 slides that were imaged with active search and with grid search. The test set was used only to assess the final performance. It was never used for training.

3.4.2 Data splitting

In the initial explorations, we used a 4-fold cross validation to train models and obtain performance estimates (**Fig. A8** in appendix). For a training run, the training set and 3 folds of validation set were pooled while the fourth validation fold was left out. Images from the same sample have similar properties and are not independent. Therefore, the split was stratified by sample so that images from a particular sample were either all used for training or all for validation. For training of the final models, which were included in the final prediction module, the training and validation sets were pooled. In this way we use the rich data from the training set and all the representative data of the validation set. This will be referred to as the *final training set* (5121 images from 33 independent samples). The final estimation of detection performance was then obtained with the test set (2415 images from 16 independent samples).

Table 1. Taxonomic definition of the classes. *1 *Nebela tincta*, *N. pechorensis*, *N. guttata*, *N. gimlii*, *N. rotunda*, *N. bohémica*, *N. collaris*, *N. minor*, *Planocarina marginata*

Class name	Taxa	Peat indicator	Approx. Length μm (long axis)
Amphitrema_sp	<i>Amphitrema</i> sp.	strict	75
Archerella_degraded	<i>Archerella</i> sp.	strict	60
Archerella_sp	<i>Archerella</i> sp.	strict	60
Assulina_sp	<i>Assulina</i> sp.	false	45 - 100
Euglypha_sp	<i>Euglypha</i> sp.	false	50 - 100
Heleopera_sphagni	<i>Heleopera sphagni</i>	strict	110 - 130
Hyalosphenia_elegans_aggr	<i>H. elegans</i> , <i>H. insecta</i>	strong	80 - 110
Hyalosphenia_papilio	<i>Hyalosphenia papilio</i>	strict	110 - 140
Nebela_combined	All taxa listed in *1	strong	70 - 150
Planocarina_carinata	<i>Planocarina carinata</i>	strict	150-240

Table 2. Overview of all samples and images. Comm: Commercial; L: Left; C: Center; R: Right

Role	Type	Source	Nb	Nb imgs	Nb imgs	Nb imgs	Nb imgs
			samples	grid	active L	active C	active R
Training	Peat	Direct	16			1666	
Validation	Peat	Comm.	11	1705	283	283	283
Validation	Non-peat	Comm.	6	901			
Test	Peat	Comm.	12	1200	272	272	271
Test	Non-peat	Comm.	4	400			

Table 3. Detailed overview of the number of manually annotated individuals per class

	Training	Validation		Test	
		active peat (center)	grid peat	active peat (center)	grid peat
Nb samples	16	11	11	12	12
Nb slides per sample	.	2	3	2	2
Tot slides	.	22	33	24	24
Nb images	1666	283	1705	272	1200
Amphitrema_sp	80	15	5	19	5
Archerella_degraded	0	57	17	43	2
Archerella_sp	145	109	45	149	17
Assulina_sp	263	22	7	19	0
Euglypha_sp	405	8	1	1	0
Heleopera_sphagni	36	3	0	0	0
Hyalosphenia_elegans_aggr	121	1	0	4	0
Hyalosphenia_papilio	201	6	1	7	0
Nebela_combined	221	0	0	0	0
Planocarina_carinata	27	0	0	0	0

3.5 Model training

3.5.1 Preparation of training data

In the training and validation sets the classes were unbalanced (**Tables 3**). For instance, we note that *Amphitrema* was underrepresented compared to *Archerella*. Prior to training, the class balance was improved by making multiple copies of images from less frequent classes. This was a static transform performed prior to training the actual models. This improves learning by the network because all classes are seen approximately the same number of times during training and it reduces the risk that the network is biased towards the more frequent classes. Then, static data augmentation was applied to all images of a training set, namely: random crop from 0 to 20 pixels, random rotation -3 to 3 degrees, random mixing with EMDS7 images weighted from 0.0 to 0.5, application of a small amount of elastic transform (local distortion). This adds some diversity to multiple copies of the same image. Note that training-time data-augmentation was also used (see below). Neither the left-out validation sets nor the test set were modified.

3.5.2 Models and training sessions

We chose YOLOv8 models. In one training session, the *final training set* was used to train one model. We performed several replications for each session with different random initialization of model weights. This allowed to assess the between session variability and serves as a basis for ensemble prediction. During training, images were resized to a width of 512 pixels. Training-time data augmentation was applied. The defaults augmentation of YOLOv8 are designed for photo images and performed poorly with our images. Thus, we increased random rotation to a range between -180 and +180 degrees, reduced random re-scaling to 0.2, and increased mix-up probability to 0.5. Models were trained for 500 epochs and we used the last value of the weights for prediction (no early stopping). Label smoothing was applied. All 10 classes (**Table 1**) were used for training. A well-established principle of ensembles is that the base predictors should be as diverse as possible (for a recent review see [61]). Therefore, the random seed was different in each session: this leads to different values of the initial weights (except for pretrained models) and different realizations of the random data augmentation. This is expected to give trained models that behave differently on individual items but similarly on average. We assessed four model sizes (NANO, SMALL, MEDIUM, LARGE) crossed with two weight initialization procedures (random initialization vs. weights pretrained with COCO dataset).

3.6 Performance evaluation

Performance was assessed exclusively with the test set by comparing the true bounding boxes with the predicted boxes. An Intersection over Union (IoU) value >0.5 was used to declare a positive detection. Precision-recall curves were constructed by incorporating the confidence value. The Average Precision (AP) was then obtained with the trapezoidal rule. PR-curves and AP were computed for each class separately, but only for classes with sufficient validation or test data.

3.7 Final prediction module via ensemble

Based on the model comparison (**Fig. A9** in appendix), final models were from YOLOv8 MEDIUM trained with random initialization of the weights. In the final prediction module, we used eight models trained in independent sessions with the *final training set* (pooled training and validation data). Each model predicts boxes and confidence values for each class. When the confidence from individual models was < 0.01 the prediction was suppressed to avoid cluttering the process with many irrelevant boxes. Note that predicted boxes from different models that detected the same object are never perfectly aligned and there is no in-built way to identify them as belonging to the same object. Therefore, we developed a post processing method to associate several predictions that detected the same object. First, the Intersection over Union (IoU) of all predicted boxes of the same class were computed for all pairs of boxes in each image. A between-box distance metric was defined as $D = 1 - \text{IoU}$ such that perfectly overlapped boxes have $D=0$ while non-overlapped boxes have $D=1$. Second, predicted

boxes that were sufficiently close on the image were aggregated via agglomerative hierarchical clustering (AHC) using D as the distance metric. This was done separately for each class and image. Therefore, only a small number of boxes at a time were processed with AHC. Confidence values of 0.0 were imputed for missing predictions (i.e. less than eight predictions associated to an object). The value of 0.0 is the lowest possible and a natural choice when an object was not detected at all. Finally, the mean of the eight predicted confidence values (incl. imputed zeros) was used as ensemble score and the corresponding crop was extracted from the prediction with max confidence value. To allow efficient decision by a human expert, the crops were plotted on a single summary image ranked by their ensemble score (highest value first, left to right and then top to bottom). Additionally, the ensemble score was color-coded with a fixed mapping such that high scores are in blue and low scores in red. In this way low score crops can be easily spotted. Examples of summary images are shown in **Fig. 10, 11, 13** and **15**.

4 Results

4.1 Testate amoebae in commercial peat samples

For the descriptive analysis of commercial peat samples (this subsection), the separation between the validation and test sets was not needed; the two sets were therefore pooled. The training set was not included because it was obtained from samples directly extracted from peatland by scientific personnel which were not commercially processed.

Archerella active: With active search, where the full slides were scrutinized, well-preserved *Archerella* sp shells were found in 22 out of 23 commercial peat samples from the pooled validation and test set (**Fig. 7**). The between-sample variability was considerable, with *Archerella* sp counts ranging from 0 to 23. Overall, 258 shells (Validation 109, Test 149) in good state of *Archerella* sp were found in 46 slides with active search (Validation 22, Test 24 slides). From this we can estimate a frequency of 5.61 shells per slide in commercial samples (**Table 4**).

Archerella grid: With grid search, well-preserved *Archerella* sp shells were found in 17 out of 23 samples from the pooled validation and test set (**Fig. 7**). Note that with the manual grid search only a small part of the slides was scrutinized (**Fig. 4 A**) which explains the smaller counts. Overall, 62 well-preserved *Archerella* sp shells (Validation 45, Test 17) were found in 2902 grid search images (Validation 1705, Test 1200 images). From this we can estimate a frequency of 0.021 shells per image (**Table 5**) when automated scanning will be applied to commercial samples. *Archerella* was by far the most frequent testate amoeba taxon found in commercial samples. Masked or degraded shells (*Archerella* degraded) were less frequently found than well-preserved *Archerella* sp. shells.

Other frequently observed taxa active: Two other genera, *Amphitrema* sp. and *Assulina* sp., were consistently found in many commercial samples but both were clearly less frequent than *Archerella*. With active search, shells from *Amphitrema* were found in 13 out of 23 and shells from *Assulina* in 19 out of 23 commercial samples from the pooled validation and test set (**Fig. 7**). For both species the counts were low with 0 to 5 individuals per sample. For *Amphitrema* and *Assulina* we estimate a frequency in commercial samples of 0.74 and 0.89 shells per slide respectively (**Table 4**).

Other frequently observed taxa grid: With grid search, shells from *Amphitrema* or *Assulina* were found but less often (**Fig. 7**). The grid size of 100 or 150 image was clearly insufficient to catch enough shells of these two species. However, with the higher acquisition rate of automated microscopes, higher counts can be expected.

Rarer taxa: A total of 13 shells of *Hyalosphenia papillo*, a strict peat indicator, were found overall with active search. Again, with the higher acquisition rate of automated microscopes,

higher counts can be expected. Finally, shells of the five other species were rarely found in commercial samples (**Table 4**, bottom rows).

4.2 Testate amoebae in commercial non-peat samples

All 10 samples (1301 images, see **Tables 2, A3, A4**) of peat-free substrates and components were searched for testate amoebae. Only a small number of testate amoeba shells were found and they were generally degraded or masked and thus more difficult to identify as compared to most shells found in the peat samples. The following potential testate amoeba morpho-taxa were found: *Centropyxis aerophila*-type (N=10), *Euglypha* sp. (N=1), *Assulina* sp. (N=1), and *Diffflugia lucida*-type (N=1). However, these identifications are partly tentative as the shells were either partly degraded or not clearly visible in the images (e.g., due to their position or masked by other particles). This result illustrates how using an automated microscope to increase the number of images will improve the assessment of the testate amoeba diversity and abundance in samples.

Table 4. Active search: Testate amoeba shell counts from all commercial peat sample (validation and test set) and resulting estimates of the number of shells per slide for each class.

Active search peat (center only)	Validation	Test	Shells per slide
Nb samples	11	12	.
Nb slides per sample	2	2	.
Total nb slides	22	24	46
Amphitrema_sp	15	19	0.739
Archerella_degraded	57	43	2.174
Archerella_sp	109	149	5.609
Assulina_sp	22	19	0.891
Euglypha_sp	8	1	0.196
Heleopera_sphagni	3	0	0.065
Hyalos_elegans_aggr	1	4	0.109
Hyalosphenia_papilio	6	7	0.283
Nebela_combined	0	0	0.000
Planocarina_carinata	0	0	0.000

Table 5. Grid search: Testate amoeba shell counts from all commercial peat sample (validation and test set) and resulting estimates of the number of shells per image for each class.

Grid search peat	Validation	Test	Shells per image
Nb images	1705	1200	2905
Amphitrema_sp	5	5	0.003
Archerella_degraded	17	2	0.007
Archerella_sp	45	17	0.021
Assulina_sp	7	0	0.002
Euglypha_sp	1	0	0.000
Heleopera_sphagni	0	0	0.000
Hyalos_elegans_aggr	0	0	0.000
Hyalosphenia_papilio	1	0	0.000
Nebela_combined	0	0	0.000
Planocarina_carinata	0	0	0.000

	Active search (center)				Grid search			
	Amphitrema_sp	Archerella_degraded	Archerella_sp	Assulina_sp	Amphitrema_sp	Archerella_degraded	Archerella_sp	Assulina_sp
Validation	0	4	4	0	1	0	1	0
	4	6	13	4	3	6	14	1
	0	6	11	2	0	2	4	1
	0	3	14	2	0	3	2	0
	1	2	21	2	1	3	4	2
	3	16	18	4	0	1	7	0
	1	6	4	1	0	1	2	0
	0	0	4	1	0	0	2	0
	5	8	10	3	0	1	5	0
	1	6	10	0	0	0	4	0
	0	0	0	3	0	0	0	3
Test	3	3	16	1	1	1	1	0
	4	6	14	4	0	0	1	0
	0	1	4	1	0	0	0	0
	0	0	6	1	0	0	0	0
	1	5	15	0	0	0	2	0
	0	1	6	2	0	0	0	0
	5	7	11	1	2	1	0	0
	0	11	23	3	0	0	2	0
	1	4	23	2	1	0	4	0
	3	4	13	1	0	0	4	0
	0	0	1	0	0	0	0	0
	2	1	17	3	1	0	3	0

Figure 7. Matrix view of testate amoeba distribution of all commercial peat samples (one sample per row, validation: 11 samples, test 12 samples). Values are the count of individual shells found with active and grid search. Validation set: Grid and active search images from different slides. Test set: Grid and active search images from same slides.

4.3 Model comparison from cross validation

For comparison of model sizes and training procedures the validation data was used. Only four classes had a count large enough to estimate the average precision (AP) and were reported. We found that the larger models MEDIUM and LARGE performed better than the smaller models NANO and SMALL (supplement). For the classes *Archerella* and *Amphitrema* (both peat specific), we found that the models initialized with random weights performed equal or better than the models initialized with pre-trained weights (**Fig. A9** in appendix). Notably for *Archerella*, size and training procedure had virtually no impact on performance, with all models resulting in AP around 0.80 (**Fig. A9** in appendix). Based on these findings, the MEDIUM model initialized with random weights was chosen as final model and used in the final prediction module.

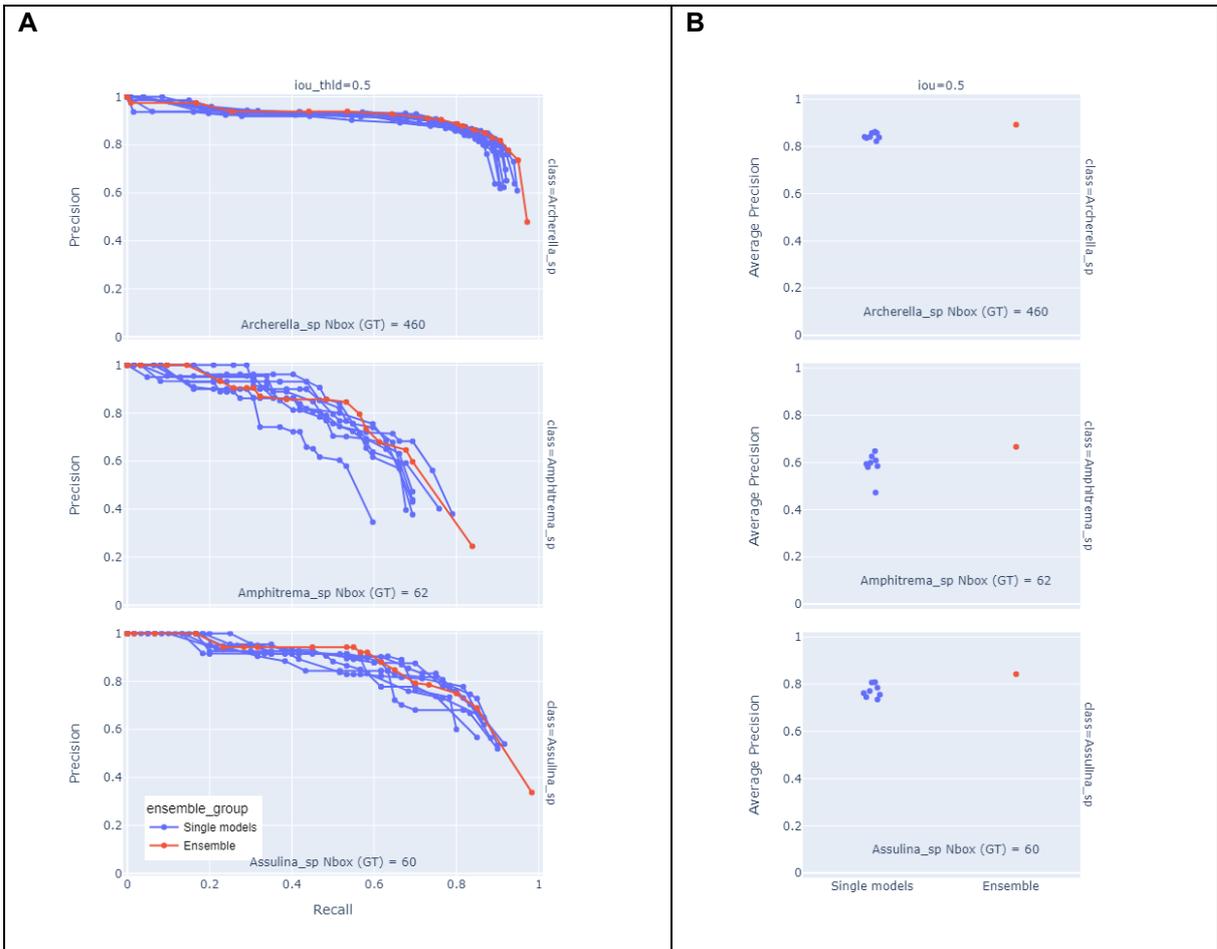


Figure 8. (A) Precision-recall curves obtained by predicting all test images with the 8 models individually (blue lines) and the ensemble (red line). Test data from grid and active search pooled (2415 images). Nbox (GT): number of annotated ground truth boxes. **(B)** Average Precision (AP) obtained from the curves on the left with the 8 models individually (blue dots) and the ensemble (red dots)

4.4 Performance with test data

The final models (MEDIUM, random initialization of weights) were trained with the *final training set* (pooled training and validation data; 5121 images). Then, performance estimates were obtained by predicting the complete test data (2415 images: grid search and active search C, L, R). See **Table 2** for details. The estimates for the three classes with sufficient instance in the test data are given in **Fig. 8 B**. The performance of individual models was high with APs above 0.8 for *Archerella*, around 0.6 for *Amphitrema*, and around 0.8 for *Assulina*. The performance of individual models was stable over the eight sessions. Noticeably, the behaviour of the score, represented by the shape of the PR curves, was also stable across individual models (**Fig. 8 A**). The performance of the ensemble model was higher than the individual models for the three classes. The AP metric is valuable for ML experts to compare several models. However, it is not very telling in terms of practical usefulness because prediction successes and errors are summarized in a single number. Therefore, a practical illustration of the performance obtained with the test data is given below. In this illustration correct and wrong predictions are reported in detail and this gives a concrete sense of the practical usefulness.

4.5 Illustration of module with test data

This section shows the summaries that were automatically generated by the prediction module together with estimates of the detection performance given as precision and recall. It is meant to illustrate a possible practical application of the module. More conventional performance metric obtained from the test set are reported in the previous section. The whole process, including prediction by an ensemble of eight models and post processing, took 581

seconds for 1872 images on a GPU accelerated desktop computer (HP Z4 workstation with NVIDIA-A6000 GPU). This corresponds to 0.31 seconds per image; thus for 1000 images it would take approximately 5 minutes.

4.5.1 *Archerella* grid only

In the first evaluation, 1600 test images, only from grid search, were passed through the prediction module. In these images the frequency of occurrence of *Archerella* shells was low at 1% (17 individuals in 1600 images). The 17 true *Archerella* shells are shown in **Fig. 9**. We consider this to be a realistic scenario because the images result from a grid search. It also represents a difficult challenge with only 17 targets hidden in 1600 images containing many objects of different shapes and sizes (mostly more or less degraded plant remains). The prediction modules detected a total of 208 crops (*Archerella* candidates).

With the usual detection threshold of 0.5, 71% (12 out of 17) individuals were detected (Recall: $12/17 = 0.71$). All 12 detected crops were correct detection (Precision: $12/12 = 1.00$). Using the lowest detection threshold of 0.0, 88% (15 out of 17) individuals were detected (Recall: $15/17 = 0.88$) at the cost of a lower precision of 0.07. The first row of the summary (12 crops) contains exclusively true *Archerella* (**Fig. 10**). Looking at the second to fourth rows, we found that three more crops were true *Archerella*. All other crops from the fifth line to the bottom were not *Archerella* and most had an ensemble score close to 0. See supplementary file for full image.

4.5.2 *Archerella* grid and active

In the second evaluation, the grid and active search (centre only) test data, were passed through the prediction module. The frequency of occurrence of individuals per image of *Archerella* was higher with 9% (166 individuals in 1872 images). The prediction modules detected a total of 450 crops (candidates), each with a value of the ensemble score on which a decision threshold can be applied. Using a detection threshold of 0.5 on the ensemble score, we obtained a recall of 0.76 and a precision of 0.91. Using the lowest detection threshold of 0.0 resulted in a recall of 0.96 at the cost of a lower precision of 0.35. Looking at the 6 first rows of the summary (**Fig. 11**), we see that all top-ranked crops clearly belong to *Archerella*. Note that two items marked with a yellow frame were labelled as *Archerella* degraded because they were partially masked. For all practical purpose these are not false positives. See supplementary file for full image.

4.5.3 *Amphitrema* grid and active

The frequency of occurrence of *Amphitrema* shells per image was low with only 1.3% of images containing an *Amphitrema* (24 individuals in 1872 images) which represents a difficult detection challenge. The 24 true *Amphitrema* shells are shown in **Fig. 12**. The prediction modules detected a total of 135 crops (candidates). Using a detection threshold of 0.5, we obtained a recall of 0.33 and a precision of 0.88. Using the lowest detection threshold of 0.0 we obtained a recall of 0.79 at the cost of a lower precision of 0.14. In the first row of the summary, 8 out of 9 crops were true *Amphitrema* and looking at the three top rows, we found that 16 out of 28 crops were true *Amphitrema* (**Fig. 13**). See supplementary file for full image.

4.5.4 *Assulina* grid and active

The frequency of occurrence of *Assulina* per image was low with only 1.0% of images containing a shell (19 individuals in 1872 images), which represents a difficult detection challenge. The 19 true *Assulina* shells are shown in **Fig. 14**. The prediction modules detected a total of 92 crops (*Assulina* candidates). Using a detection threshold of 0.5, we obtained a recall of 0.79 and a precision of 0.94. Using the lowest detection threshold of 0.0 we obtained recall of 1.00 at the cost of a lower precision of 0.21. In the first row of the summary, 8 out of 11 crops were true *Assulina* and looking at the first three rows, we found that 17 out of 34 crops were true *Assulina* (**Fig. 15**). See supplementary file for full image.

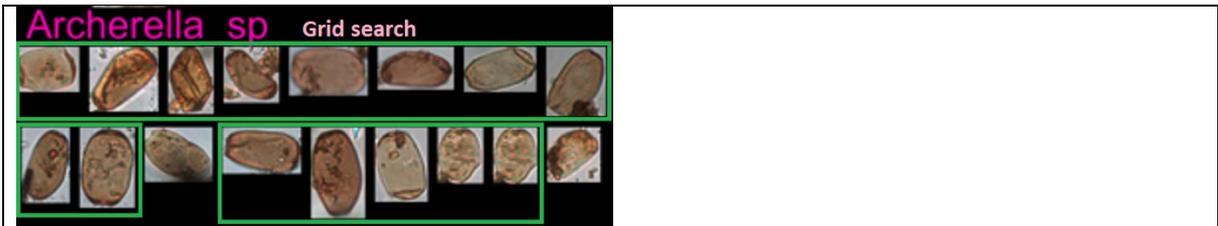


Figure 9. Ground truth in test set for N=17 *Archerella* sp from grid search. The order is arbitrary. The 15 individual that were detected (**Fig. 10**) are framed in green.



Figure 10. Automatically detected candidates of *Archerella* sp ranked by ensemble score (only top rows shown) obtained from 1600 grid search images of test set. The correctly detected shells are framed in green.



Figure 11. Automatically detected candidates of *Archerella* sp ranked by ensemble score (only top rows shown) obtained from 1600 grid and 272 active search images of test set. Two items that were labelled as *Archerella* degraded due to masking are framed in yellow; the remaining 73 crops are correctly detected shells.

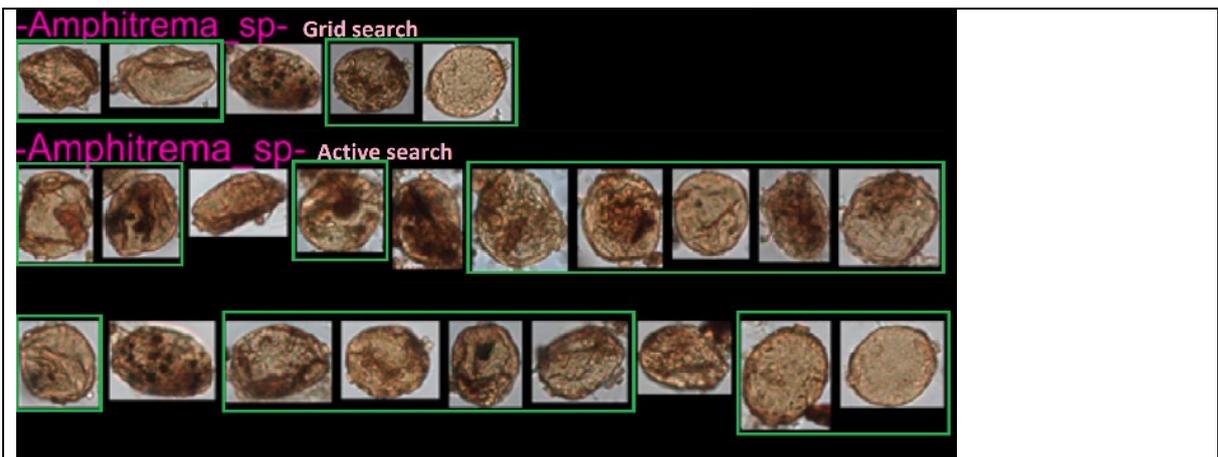


Figure 12. Ground truth in test set for N=24 *Amphitrema* sp from grid search (top) and active search (bottom). The order is arbitrary. The 19 individual that were detected (**Fig. 13**) are framed in green.

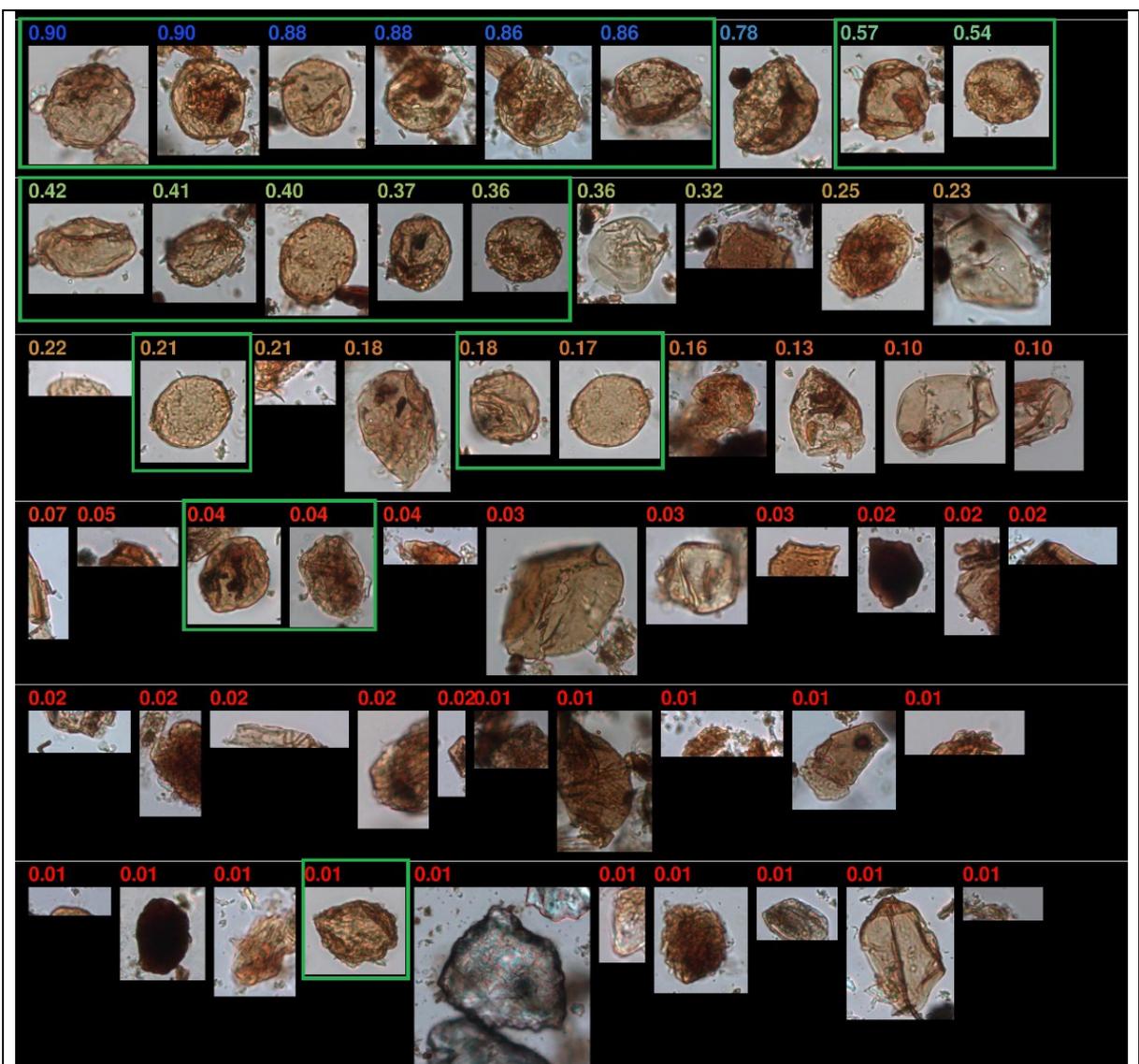


Figure 13. Automatically detected candidates of *Amphitrema* sp ranked by ensemble score (only top rows shown) obtained from 1600 grid and 272 active search images of test set. The correctly detected crops are framed in green.

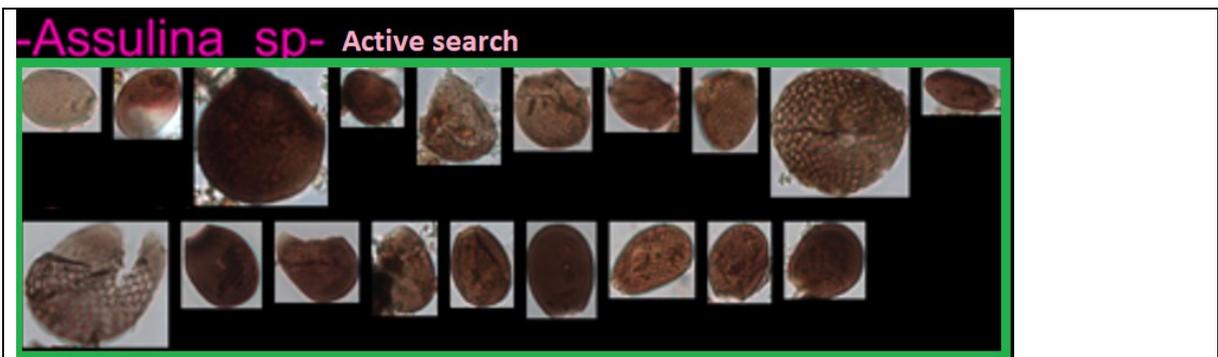


Figure 14. Ground truth in test set for N=19 *Assulina* sp active search (zero individuals were found with grid search). The order is arbitrary. All 19 individual were detected (**Fig. 15**) and are framed in green.



Figure 15. Automatically detected candidates of *Assulina* sp ranked by ensemble score (only top 5 rows shown) obtained from 1600 grid and 272 active search images of test set. The correct detections are framed in green.

5 Discussion

5.1 Shells preserved in commercial peat products

The shells of several peat-specific testate amoeba taxa were present in commercially processed substrates containing only peat. For two of the smaller peat-specific taxa (*Archerella* and *Amphitrema*), well-preserved specimens were found in many samples. These two species can therefore be used as indicators of peat presence in commercially processed products. Another taxon (*Assulina*) which is common in peatlands but also frequent in acidic soils outside of peatlands (e.g. coniferous forest) was also frequently found in commercial substrates. This further confirms the resistance of small testate amoeba shells to industrial processes. This observation is not surprising as *Assulina* is among the most resistant testate amoeba taxon to degradation; it is not even destroyed by hydrofluoric acid treatment used for pollen preparation [65]. While a few shells of larger species were found (*Hyalosphenia* sp), they were infrequent and often degraded. For this reason, they seem less optimal as indicators of peat in commercial products. However, if a larger number of images can be automatically acquired, we can expect to find enough shells in good state from *Hyalosphenia* sp to be used as indicators. Based on the above findings, we are confident that the automated detection of peat-specific testate amoebae is a valid method to detect peat in substrate. The results from *Archerella* and *Amphitrema* have already been turned into an actionable solution presented here.

5.2 Commercial peat-free substrates and components

A total of 1301 images from commercial media labelled as peat-free was carefully checked. Several specimens of *Centropyxis aerophila*-type and one *Diffulgia lucida*-type were found. Although these species can be found in peatlands, they are also frequent in other habitats including caves [66]. The samples also included many fungal spores and some conifer pollen. The fact that several specimens of testate amoeba species were found is positive as it shows that they can be found in such samples. The absence of the typical peat indicators can be interpreted as proof that the samples were indeed peat-free. In addition, the absence of *Assulina* further demonstrates that the samples were not even from acidic litter collected e.g. in coniferous forests. If the substrate is not acidic this taxon would indeed not be expected to be found as it is clearly an indicator of low pH [67].

5.3 Strengths and limitations

Our method is strong to demonstrate the presence of peat. That is, if shells of peat-specific testate amoebae were found, this represents strong evidence that peat is in the substrate. In addition, the method automatically provides supporting evidence in the form of crops, which can be interpreted by human experts (Fig. 10, 11, 13, 15 and images in supplement). There is a between-sample variability in the frequency of peat-specific amoebae. It remains possible that the species could be absent in peat from certain provenances. In this case our methods will naturally not work. The method is thus weaker to rule-out the presence of peat. That is, if peat-specific shells were not found, the substrate could still contain peat. This possible limitation however needs to be further evaluated.

Nowadays growing media distributed commercially may contain recycled peat. Such substrates are less problematic from a sustainability point of view. We noted that two indicator species are robust to commercial processing. These shells might also be conserved through the recycling process and thus our method might detect them. Fundamentally, this would not be a limitation of the method per-se but rather a consequence of its good detection performance. Note also that any detection method for peat will probably suffer the same limitation. Again, this possibility should be tested, ideally experimentally.

6 Outlook

6.1 Large-scale analysis with an automated microscope

In commercial samples only 2% of images contained an *Archerella* shell (**Table 5**). When present, the shell covered much less than 1% of the image and was often embedded in residues (example in **Fig. 5**). To obtain a clear statement regarding peat presence in each sample, we propose that at least 1000 images per sample should be searched (thus we expect to catch approximately 20 *Archerella* shells in pure peat). For mixed substrate we would need even larger numbers of images to capture approximately 20 *Archerella*. Performing this manually is simply not realistic for a large-scale analysis with hundreds of samples. Our proposed decision support system automates the tedious detection and extraction process and produces a concise summary as a ranked list of *Archerella* candidates where the most promising items are shown first. The final step, manual confirmation and counting by a human operator, is expected to take at most a few minutes per sample. The proposed decision support process significantly reduces the amount of human effort required.

Our decision support system is an important step to enable the large-scale monitoring of commercial substrates from multiple sources. Let's assume for instance a scenario with 100 samples, each with 1000 images. We would need around 8.3 hours to process all 100'000 images on a single GPU accelerated computer. In addition, the process is reproducible and reliable because the final prediction module incorporates the expertise of very specialized domain experts (CV, EM, CD) via highly curated datasets. This would allow the decision support process to be safely performed by less experienced staff who underwent a short training. The process is also transparent because the generated crops can be stored for each sample as evidence and allow independent review by third parties if needed.

6.2 Need for automated microscope

The decision support system requires many images per sample that must be acquired by systematically scanning all the material under the slide. In this study we did this manually, but for a large-scale analysis as proposed above this is simply not feasible. Thus, an automated scanning microscope will be needed. However, note that this would only accelerate the image acquisition per-se but not the preliminary lab work to obtain sample material, process the samples, and prepare the microscopy slides.

6.3 Development of a method to quantify peat percentage

Horticultural substrate often contains a mix of peat and other substrates (e.g. compost). It would be useful to estimate the proportion of peat in such mixtures. The present results (**Table 4** and **Fig. 7**) show that *Archerella* was consistently present in natural peat from Europe. Another strict peat indicator, *Amphitrema* was found in half of the peat samples despite the small number of images available per sample. If larger number of images per samples can be gathered via automated microscopy techniques, we can expect to find sufficient shells of peat indicator taxa to estimate the shell density in each sample accurately: e.g., number of shells per mg substrate. Provided that the shell density in pure peat and its natural variation is known, this could be used to estimate the proportion of peat in substrate mixtures. The natural variation will obviously add uncertainty to this estimate. The results obtained here with a rather small number of images are promising. Thus, as a next step, it would be interesting to assess the variation in natural frequency, especially of *Archerella* and *Amphitrema*, in peat extracted from many European provenances.

6.4 Cultivated *Sphagnum* moss

It would be interesting to test our method on *Sphagnum* moss issued from *Sphagnum* farming. This type of farming describes the cultivation of *Sphagnum* moss on wet soils (often degraded

peat soils that are rewetted). *Sphagnum* can be harvested every few years and is used as peat-alternative in substrates. As the *Sphagnum* is only cut superficially, the wet soils can act as important carbon storages. *Sphagnum* farming is being developed and commercial production exists in some countries such as Germany and Chile. It is also considered as a sustainable alternative to peat extraction in Europe. Germany for example promotes *Sphagnum* farming on degraded peat soils [73]. As cultivated *Sphagnum* moss is seen as a sustainable alternative to peat, our method should ideally be able to distinguish these two products in growing media.

6.5 Geographical scope of the method

The present study was based on samples from European provenances. We have shown that *Archerella* and *Amphitrema* were frequently found in these samples. We are therefore confident that our method will be applicable for sample originating from Europe. The species composition and morphological features of testate amoeba are probably different for other continents and especially for region in the southern-hemisphere. Therefore, we do not expect the current version of our method to be applicable for samples originating from outside of Europe. Fortunately, the data and models could be extended with peat-indicator species from other continents.

6.6 Extension to other applications

The present results show that deep neural networks could be adapted to effectively detect soil microorganisms in a very specific setting. The task was challenging due to the degraded state of the amoeba shells in commercial products and to their low frequency of occurrence. Adding to the difficulty, the shells were embedded in plant residues and particles with diverse shapes and textures. This is likely to be the case in similar applications involving images from natural soil samples or products derived thereof. Despite these complications, well performing models could be trained without dependence on pre-trained models and with a rather small development dataset of a few thousand images. A big workload was the manual acquisition of images. Fortunately, in future applications, this can be achieved with automated microscopy devices, which is expected to considerably reduce the human workload. The data annotation (bounding boxes) could be done efficiently thanks to the availability of user-friendly tools that can be used collaboratively via the web. We have also shown that a few thousand images together with a carefully tuned data augmentation strategy are sufficient to obtain actionable and useful models. We believe that this will translate to future projects and thus it will keep the annotation burden within reasonable boundaries. Finally, model selection and fine-tuning of the training procedure was time consuming. Fortunately, the developed procedure can be used for new classes, thus this work must not be repeated. Based on the above insights, the extension of the model to more microorganism classes seems a feasible endeavour. For instance, we could make models that detect many taxa from direct soil probes. This would open new possibilities for large scale assessments of soil biodiversity or the continuous monitoring of soil health at restoration sites.

7 Acknowledgments

We would like to sincerely thanks the following companies and institutions that have provided substrate samples for our study: Gramoflor GmbH & Co. KG, Klasmann-Deilmann GmbH, ökohum gmbh, Stender GmbH, RICOTER Erdaufbereitung AG, and Zürcher Hochschule für Angewandte Wissenschaften (ZHAW). We greatly appreciate the high-quality material that they have supplied. These materials have significantly contributed to the success of our project.

8 Conclusions

- Peat-specific testate amoebae were recovered from commercial substrate containing peat. Two taxa, *Archerella* and *Amphitrema*, are robust indicators of peat presence in such products.
- Deep neural networks were successfully trained and tested with a small application-specific data set. This illustrates the maturity of these algorithms for real-world applications.
- We propose a decision process where large image collections are automatically batch processed and human experts can quickly review a ranked list of small crops to take the final decision.
- Our method renders feasible large-scale monitoring of peat presence in commercial substrate from multiple sources.
- An automated image acquisition procedure will be needed to render the method practical.

9 References

- [1] Steve Froking et al. "Peatlands in the Earth's 21st century climate system". In: *Environmental Reviews* 19.NA (2011), pp. 371–396.
- [2] Jiren Xu, Paul J Morris, Junguo Liu, and Joseph Holden. "PEATMAP: Refining estimates of global peatland distribution based on a meta-analysis". In: *Catena* 160 (2018), pp. 134–140.
- [3] Jens Leifeld, Chloé Wüst-Galley, and Susan Page. "Intact and managed peatland soils as a source and sink of GHGs from 1850 to 2100". In: *Nature Climate Change* 9.12 (2019), pp. 945–947.
- [4] Intergovernmental Panel on Climate Change (IPCC). *Climate Change 2022 – Impacts, Adaptation and Vulnerability: Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 2023.
- [5] Ralph JM Temmink et al. "Wetscapes: Restoring and maintaining peatland landscapes for sustainable futures". In: *Ambio* 52.9 (2023), pp. 1519–1528.
- [6] Sate Ahmad, Haojie Liu, Anke Günther, John Couwenberg, and Bernd Lennartz. "Longterm rewetting of degraded peatlands restores hydrological buffer function". In: *Science of the Total Environment* 749 (2020), p. 141571.
- [7] Olivier Hirschler and Bernhard Osterburg. "Peat extraction, trade and use in Europe: a material flow analysis." In: *Mires & Peat* 28 (2022).
- [8] Ralf Meisterfeld. *The illustrated guide to the protozoa: Order Arcellinida Kent, 1880*. Ed. By G. F. Leedale & P. Bradbury J. J. Lee. 2002.
- [9] Ralf Meisterfeld. *The illustrated guide to the protozoa: Testate amoebae with filopodia*. Ed. By G. F. Leedale & P. Bradbury J. J. Lee. 2002.
- [10] Otto Harnisch. "Einige Daten zur rezenten und fossilen testaceen Rhizopodenfauna der Sphaggen". In: *Arch. Hydrobiol* 18.3 (1927), p. 245.
- [11] Odile Boeuf and Daniel Gilbert. "Presence of testate amoebae (genus: *Trinema*), in the Upper Pliocene, discovered in the locality of Chilhac (Haute-Loire, France)". In: *Comptes Rendus de l'Academie des Sciences Series IIA Earth and Planetary Science* 8.325 (1997), pp. 623–627.
- [12] Andrew Barber, Peter A Siver, and William Karis. "Euglyphid testate amoebae (Rhizaria: Euglyphida) from an Arctic Eocene waterbody: evidence of evolutionary stasis in plate morphology for over 40 million years". In: *Protist* 164.4 (2013), pp. 541–555.
- [13] Susannah M Porter and Andrew H Knoll. "Testate amoebae in the Neoproterozoic Era: evidence from vase-shaped microfossils in the Chuar Group, Grand Canyon". In: *Paleobiology* 26.3 (2000), pp. 360–385.
- [14] Alexander R Schmidt, Eugenio Ragazzi, Olimpia Coppellotti, and Guido Roghi. "A microworld in Triassic amber". In: *Nature* 444.7121 (2006), pp. 835–835.
- [15] Luana Morais et al. "Carbonaceous and siliceous Neoproterozoic vase-shaped microfossils (Urucum Formation, Brazil) and the question of early protistan biomineralization". In: *Journal of Paleontology* 91.3 (2017), pp. 393–406.
- [16] Edward AD Mitchell, Daniel J Charman, and Barry G Warner. "Testate amoebae analysis in ecological and paleoecological studies of wetlands: past, present and future". In: *Biodiversity and conservation* 17 (2008), pp. 2115–2137.
- [17] Graeme T Swindles et al. "Widespread drying of European peatlands in recent centuries". In: *Nature Geoscience* 12.11 (2019), pp. 922–928.
- [18] Yangmin Qin et al. "Developing a continental-scale testate amoeba hydrological transfer function for Asian peatlands". In: *Quaternary Science Reviews* 258 (2021), p. 106868.

- [19] R Timothy Patterson, Andrew Dalby, Arun Kumar, Lori A Henderson, and Robert EA Boudreau. "Arcellaceans (thecamoebians) as indicators of land-use change: settlement history of the Swan Lake area, Ontario as a case study". In: *Journal of Paleolimnology* 28 (2002), pp. 297–316.
- [20] Luiz Felipe Machado Velho, Fábio Amodêo Lansac-Tôha, and Luis Mauricio Bini. "Influence of environmental heterogeneity on the structure of testate amoebae (Protozoa, Rhizopoda) assemblages in the plankton of the upper Paran a river floodplain, Brazil". In: *International Review of Hydrobiology: A Journal Covering all Aspects of Limnology and Marine Biology* 88.2 (2003), pp. 154–166.
- [21] Zai-Chao Yang, Zhi-Hui Wang, and Zhao-Hui Zhang. "Biomonitoring of testate amoebae (protozoa) as toxic metals absorbed in aquatic bryophytes from the Hg-Tl mineralized area (China)". In: *Environmental monitoring and assessment* 176 (2011), pp. 321–329.
- [22] Nawaf A Nasser, R Timothy Patterson, Helen M Roe, Jennifer M Galloway, Hendrik Falck, and Hamed Sanei. "Use of Arcellinida (testate lobose amoebae) arsenic tolerance limits as 26 a novel tool for biomonitoring arsenic contamination in lakes". In: *Ecological Indicators* 113 (2020), p. 106177.
- [23] W Roland Gehrels, Dawn Hendon, and Dan J Charman. "Distribution of testate amoebae in salt marshes along the North American East Coast". In: *Journal of Foraminiferal Research* 36.3 (2006), pp. 201–214.
- [24] Alicia Fr sard et al. "Inferring northern peatland methane emissions from testate amoebae: A proof of concept study". In: *Mires & Peat* (2023).
- [25] Vincent EJ Jassey et al. "An unexpected role for mixotrophs in the response of peatland carbon cycling to climate warming". In: *Scientific reports* 5.1 (2015), p. 16931.
- [26] Angela L Creevy, David M Wilkinson, Roxane Andersen, and Richard J Payne. "Testate amoebae response and vegetation composition after plantation removal on a former raised bog". In: *European Journal of Protistology* 89 (2023), p. 125977.
- [27] Sinikka Jauhiainen. "Testacean amoebae in different types of mire following drainage and subsequent restoration". In: *European Journal of Protistology* 38.1 (2002), pp. 59–72.
- [28] Isabelle Koenig, Florence Schwendener, Matthieu Mulot, and Edward AD Mitchell. "Response of Sphagnum testate amoebae to drainage, subsequent re-wetting and associated changes in the moss carpet-results from a three year mesocosm experiment". In: *Acta Protozoologica* 56.3 (2017), pp. 191–210.
- [29] Fatima Laggoun-D efarge et al. "Cut-over peatland regeneration assessment using organic matter and microbial indicators (bacteria and testate amoebae)". In: *Journal of Applied Ecology* 45.2 (2008), pp. 716–727.
- [30] Graeme T Swindles et al. "Evaluating the use of dominant microbial consumers (testate amoebae) as indicators of blanket peatland restoration". In: *Ecological Indicators* 69 (2016), pp. 318–330.
- [31] Julie Valentine, Stephen R Davis, Jason R Kirby, and David M Wilkinson. "The use of testate amoebae in monitoring peatland restoration management: case studies from North West England and Ireland". In: *Acta Protozoologica* 52.3 (2013).
- [32] Emma Vickery and Daniel J Charman. "Biomonitoring of peatland restoration using testate amoebae". In: *7th INTECOL international wetlands conference, vol. Book of abstracts, Utrecht, The Netherlands. 2004*, pp. 25–30.
- [33] Valentyna Krashevskaya et al. "Micro-decomposer communities and decomposition processes in tropical lowlands as affected by land use and litter type". In: *Oecologia* 187 (2018), pp. 255–266.
- [34] Daniel Gilbert and Edward AD Mitchell. *Microbial diversity in Sphagnum peatlands. Peatlands: basin evolution and deposi-tory of records on global environmental and climatic changes. 2006*.
- [35] Milcho Todorov and Nikola Bankov. "An atlas of Sphagnum-dwelling testate amoebae in Bulgaria". In: *Advanced Books* 1 (2019), e38685.

- [36] David Singer et al. "Environmental filtering and phylogenetic clustering correlate with the distribution patterns of cryptic protist species". In: *Ecology* 99.4 (2018), pp. 904–914.
- [37] Alicja Raczkowska, Marcin Mozejko, Joanna Zambonelli, and Ewa Szczurek. "ARA: accurate, reliable and active histopathological image classification framework with Bayesian deep learning". In: *Scientific reports* 9.1 (2019), p. 14347.
- [38] Zakaria Senousy et al. "MCUa: Multi-level context and uncertainty aware dynamic deep ensemble for breast cancer histology image classification". In: *IEEE Transactions on Biomedical Engineering* 69.2 (2021), pp. 818–829.
- [39] Zakaria Senousy, Mohammed M Abdelsamea, Mona Mostafa Mohamed, and Mohamed Medhat Gaber. "3E-Net: Entropy-based elastic ensemble of deep convolutional neural networks for grading of invasive breast carcinoma histopathological microscopic images". In: *Entropy* 23.5 (2021), p. 620.
- [40] Charlotte Srykh et al. "Accurate diagnosis of lymphoma on whole-slide histopathology images using deep learning". In: *NPJ digital medicine* 3.1 (2020), p. 63.
- [41] Alessandro Ferrari, Stefano Lombardi, and Alberto Signoroni. "Bacterial colony counting with convolutional neural networks in digital microbiology imaging". In: *Pattern Recognition* 61 (2017), pp. 629–640.
- [42] Sumit Paul, Salil Batra, Khalid Mohiuddin, Mohamed Nadhmi Miladi, Divya Anand, and Osman A. Nasr. "A Novel Ensemble Weight-Assisted Yolov5-Based Deep Learning Technique for the Localization and Detection of Malaria Parasites". In: *Electronics* 11.23 (2022), p. 3999.
- [43] Fetulhak Abdurahman, Kinde Anlay Fante, and Mohammed Aliy. "Malaria parasite detection in thick blood smear microscopic images using modified YOLOV3 and YOLOV4 models". In: *BMC bioinformatics* 22 (2021), pp. 1–17. 27
- [44] Carles Rubio Maturana et al. "iMAGING: a novel automated system for malaria diagnosis by using artificial intelligence tools and a universal low-cost robotized microscope". In: *Frontiers in microbiology* 14 (2023), p. 1240936.
- [45] Padmini Krishnadas, Krishnaraj Chadaga, Niranjana Sampathila, Santhosha Rao, and Srikanth Prabhu. "Classification of malaria using object detection models". In: *Informatics*. Vol. 9. 4. MDPI. 2022, p. 76.
- [46] Pingli Ma et al. "A state-of-the-art survey of object detection techniques in microorganism image analysis: from classical methods to deep learning approaches". In: *Artificial Intelligence Review* 56.2 (2023), pp. 1627–1698.
- [47] Priya Rani, Shallu Kotwal, Jatinder Manhas, Vinod Sharma, and Sparsh Sharma. "Machine learning and deep learning based computational approaches in automatic microorganisms image recognition: methodologies, challenges, and developments". In: *Archives of Computational Methods in Engineering* 29.3 (2022), pp. 1801–1837.
- [48] Jiawei Zhang et al. "A comprehensive review of image analysis methods for microorganism counting: from classical image processing to deep learning approaches". In: *Artificial Intelligence Review* (2022), pp. 1–70.
- [49] Ran Shao, Xiao-Jun Bi, and Zheng Chen. "A novel hybrid transformer-CNN architecture for environmental microorganism classification". In: *Plos one* 17.11 (2022), e0277557.
- [50] Sergey Kosov, Kimiaki Shirahama, Chen Li, and Marcin Grzegorzec. "Environmental microorganism classification using conditional random fields and deep convolutional neural networks". In: *Pattern recognition* 77 (2018), pp. 248–261.
- [51] Jinghua Zhang et al. "LCU-Net: A novel low-cost U-Net for environmental microorganism image segmentation". In: *Pattern Recognition* 115 (2021), p. 107885.
- [52] Chih-Ming Liang, Chun-Chi Lai, Szu-Hong Wang, and Yu-Hao Lin. "Environmental microorganism classification using optimized deep learning model". In: *Environmental Science and Pollution Research* 28 (2021), pp. 31920–31932.

- [53] Marcin Dziadosz, Dariusz Majerek, and Grzegorz Lagód. "Microscopic Studies of Activated Sludge Supported by Automatic Image Analysis Based on Deep Learning Neural Networks". In: *Journal of Ecological Engineering* 25.4 (2024), pp. 360–369.
- [54] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. "You only look once: Unified, real-time object detection". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 779–788.
- [55] Peiyuan Jiang, Daji Ergu, Fangyao Liu, Ying Cai, and Bo Ma. "A Review of Yolo algorithm developments". In: *Procedia computer science* 199 (2022), pp. 1066–1073.
- [56] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. "Yolov4: Optimal speed and accuracy of object detection". In: *arXiv preprint arXiv:2004.10934* (2020).
- [57] Lesheng Sun, Ying Xu, Zhikang Rao, Juntao Chen, Zhe Liu, and Ning Lu. "YOLO algorithm for long-term tracking and detection of escherichia coli at different depths of microchannels based on microsphere positioning assistance". In: *Sensors* 22.19 (2022), p. 7454.
- [58] Abdullah, Sikandar Ali, Ziaullah Khan, Ali Hussain, Ali Athar, and Hee-Cheol Kim. "Computer vision based deep learning approach for the detection and classification of algae species using microscopic images". In: *Water* 14.14 (2022), p. 2219.
- [59] Jesús Salido, Carlos Sánchez, Jesús Ruiz-Santaquiteria, Gabriel Cristóbal, Saul Blanco, and Gloria Bueno. "A low-cost automated digital microscopy platform for automatic identification of diatoms". In: *Applied Sciences* 10.17 (2020), p. 6033.
- [60] Hechen Yang et al. "EMDS-7: Environmental microorganism image dataset seventh version for multiple object detection evaluation". In: *Frontiers in Microbiology* 14 (2023), p. 1084312.
- [61] Mudasar A Ganaie, Minghui Hu, Ashwani Kumar Malik, Muhammad Tanveer, and Ponnuthurai N Suganthan. "Ensemble deep learning: A review". In: *Engineering Applications of Artificial Intelligence* 115 (2022), p. 105151.
- [62] Gomaa, F., Kosakyan, A., Heger, T.J., Corsaro, D., Mitchell, E.A. & Lara, E. (2014) One alga to rule them all: unrelated mixotrophic testate amoebae (amoebzoa, rhizaria and stramenopiles) share the same symbiont (trebouxiophyceae). *Protist*, 165, 161-176.
- [63] Amesbury, M.J., Swindles, G.T., Bobrov, A., Charman, D.J., Holden, J., Lamentowicz, M., Mallon, G., Mazei, Y., Mitchell, E.A.D., Payne, R.J., Roland, T.P., Turner, T.E. & Warner, B.G. (2016) Development of a new pan-European testate amoeba transfer function for reconstructing peatland palaeohydrology. *Quaternary Science Reviews*, 152, 132-151.
- [64] Amesbury, M.J., Booth, R.K., Roland, T.P., Bunbury, J., Clifford, M.J., Charman, D.J., Elliot, S., Finkelstein, S., Garneau, M., Hughes, P.D.M., Lamarre, A., Loisel, J., Mackay, H., Magnan, G., Markel, E.R., Mitchell, E.A.D., Payne, R.J., Pelletier, N., Roe, H., Sullivan, M.E., Swindles, G.T., Talbot, J., van Bellen, S. & Warner, B.G. (2018) Towards a Holarctic synthesis of peatland testate amoeba ecology: Development of a new continental-scale palaeohydrological transfer function for North America and comparison to European data. *Quaternary Science Reviews*, 201, 483-500.
- [65] Payne, R.J., Lamentowicz, M., van der Knaap, W.O., van Leeuwen, J.F.N., Mitchell, E.A.D. & Mazei, Y. (2012) Testate amoebae in pollen slides. *Review of Palaeobotany and Palynology*, 173, 68-79.
- [66] Mazei, Y., Belyakova, O., Trulova, A., Guidolin, L. & Coppelotti, O. (2012) Testate amoebae communities from caves of some territories in European Russia and North-Eastern Italy. *Protistology*, 7, 42–50.
- [67] Bonnet, L. (1991) *Ecologie de quelques Euglyphidae (Thécamoebiens, Filosea) des milieux édaphiques et para édaphiques (Deuxième partie)*. *Bulletin de la Société d'Histoire Naturelle de Toulouse*, 127, 15-20.
- [68] Urech M. (2016) *Datenerhebung Torfimport und Torfverwendung in der Schweiz 2014*

- [69] BAFU (2020) Faktenblatt zur Halbierung des Torfanteils in Sackerden zwischen 2016 und 2018
- [70] Ficken, K., Barber, K. E. & Eglinton, G. (1998) Lipid biomarker, $d^{13}C$ and microfossil stratigraphy of a Scottish mountain peat bog over the last two millennia. *Organic Geochemistry*, 28, 217-237
- [71] Nichols, J. E., Booth, R. K., Jackson, S. T., Pendall, E. G. & Huang, Y. (2006) Paleohydrologic reconstruction based on n-alkane distributions in ombrotrophic peat. *Organic Geochemistry*, 37, 1505-1513
- [72] Pancost, R. D, Baas, M., van Geel, B. & Sinninghe Damsté, J, S. (2002) Biomarkers as proxies for plant inputs to peats: an example from a sub-boreal ombrotrophic bog. *Organic Geochemistry*, 33, 675-690
- [73] Gaudig, G., Krebs, M., Prager, A., Wichmann, S., Barney, M. & Capon, S. J. M.(2017) Sphagnum farming from species selection to the production of growing media: a review. *Mires and peat*, 20, 1–30
- [74] Schmilewski, G., (2008). The role of peat in assuring the quality of growing media. *Mires and Peat*.

10 Appendix

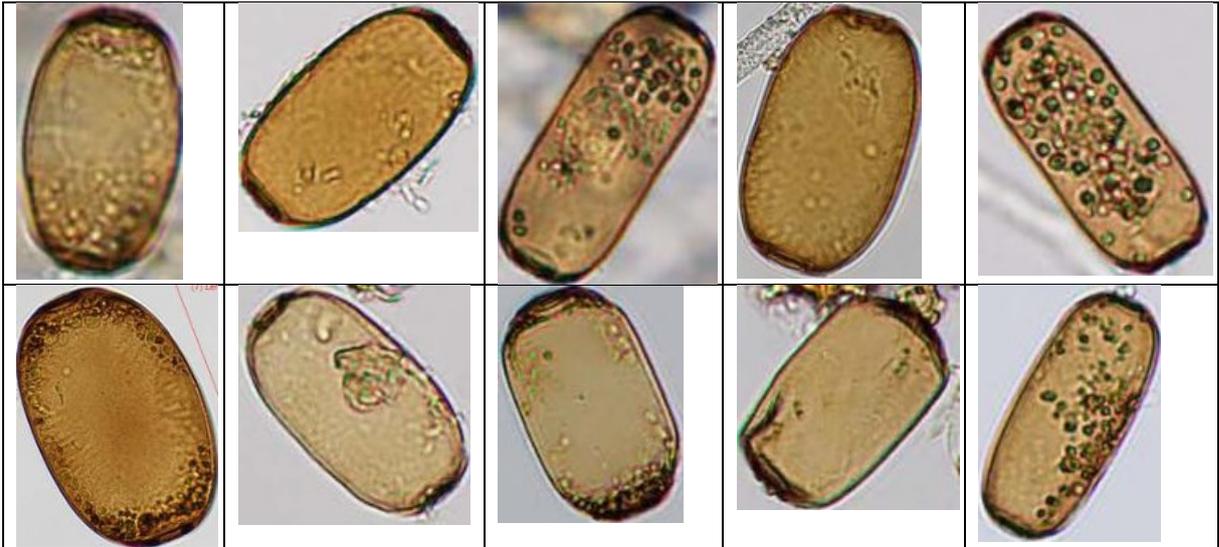


Figure A1. Shells of *Archerella* sp taken from images of direct soil probes (Training set)

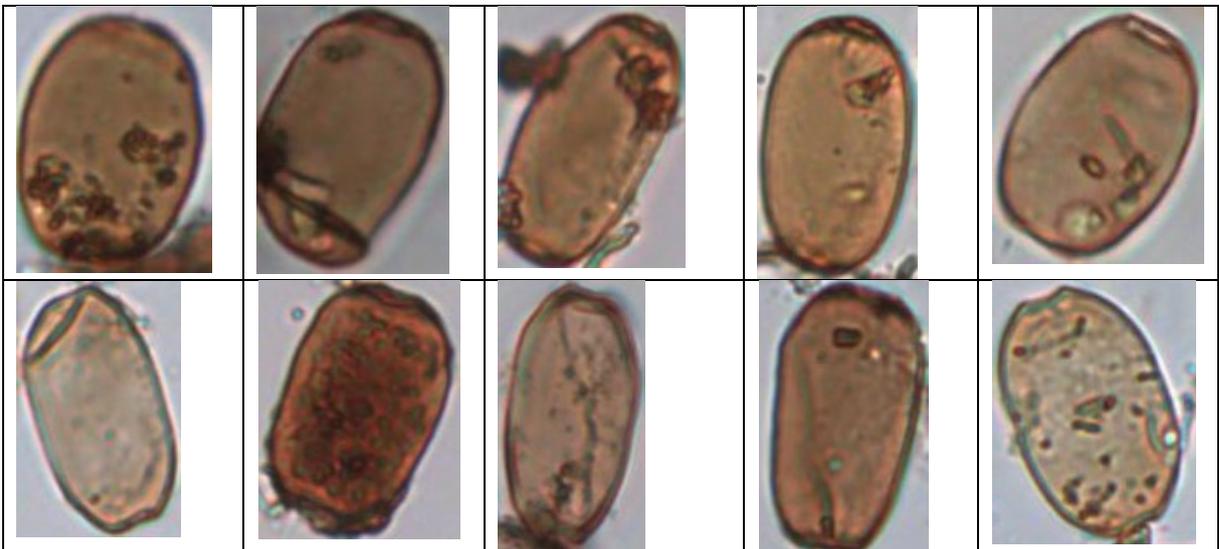


Figure A2. Shells of *Archerella* sp taken from images of commercial samples (Test set)

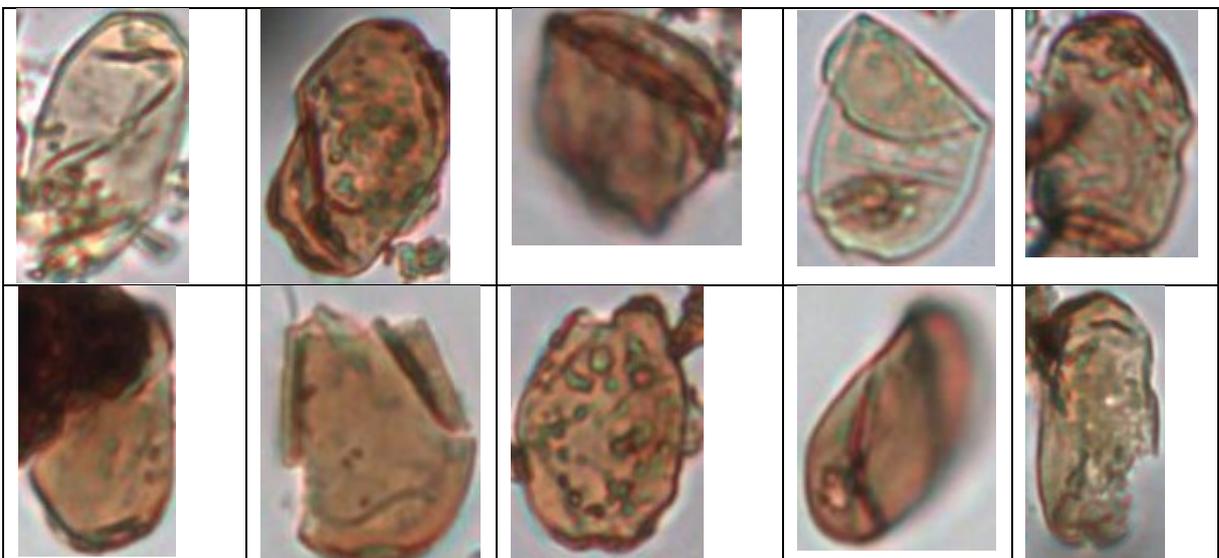


Figure A3. Shells of the class *Archerella* degraded from images of commercial samples (Test set).

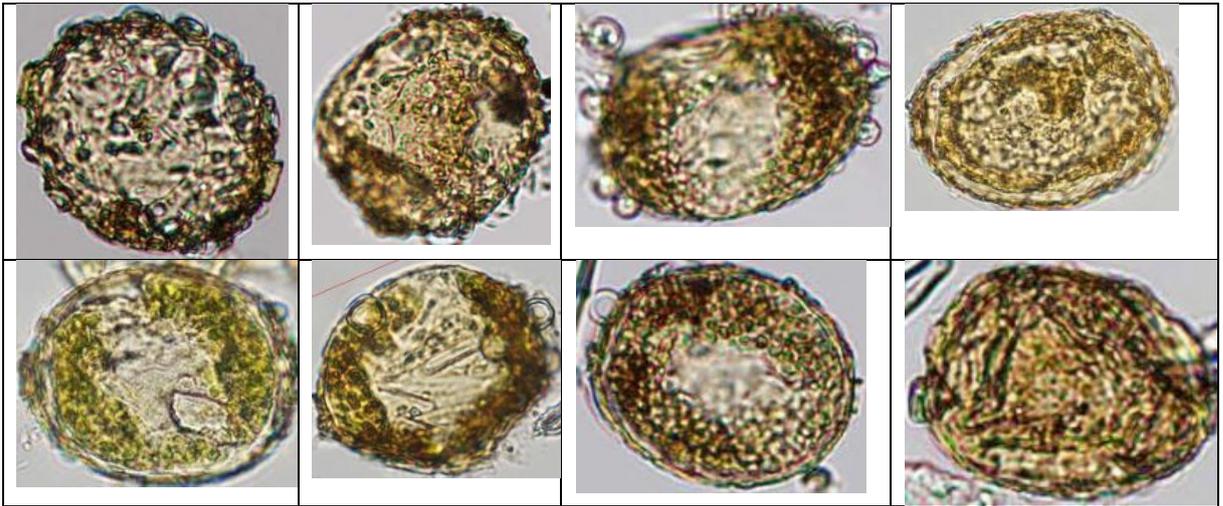


Figure A4. Shells of *Amphitrema* sp from images of direct soil probes (Training set)

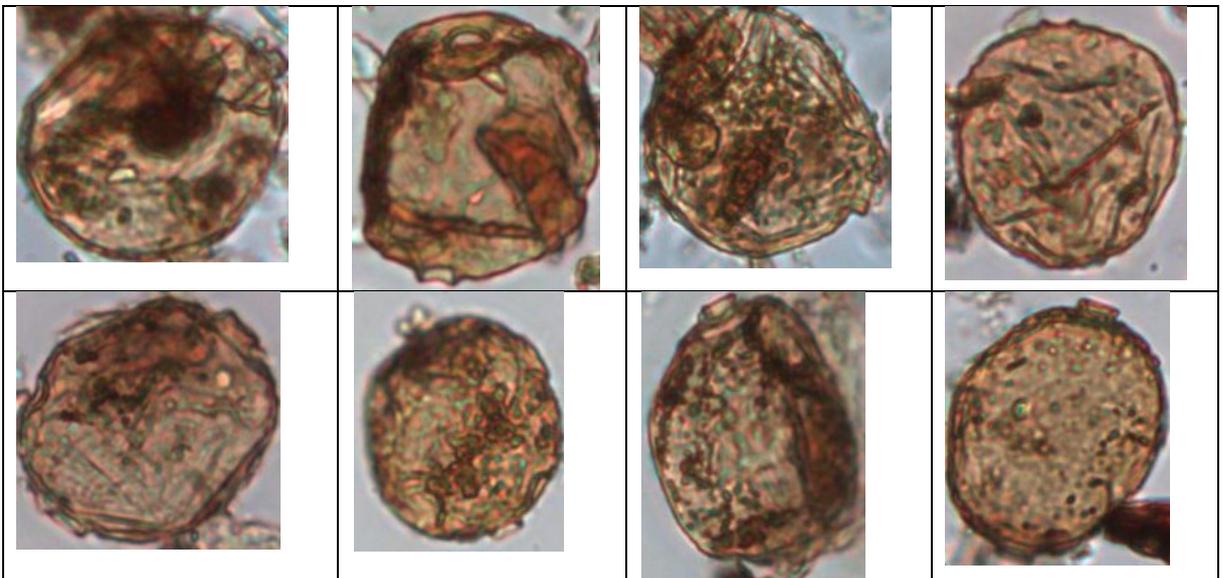
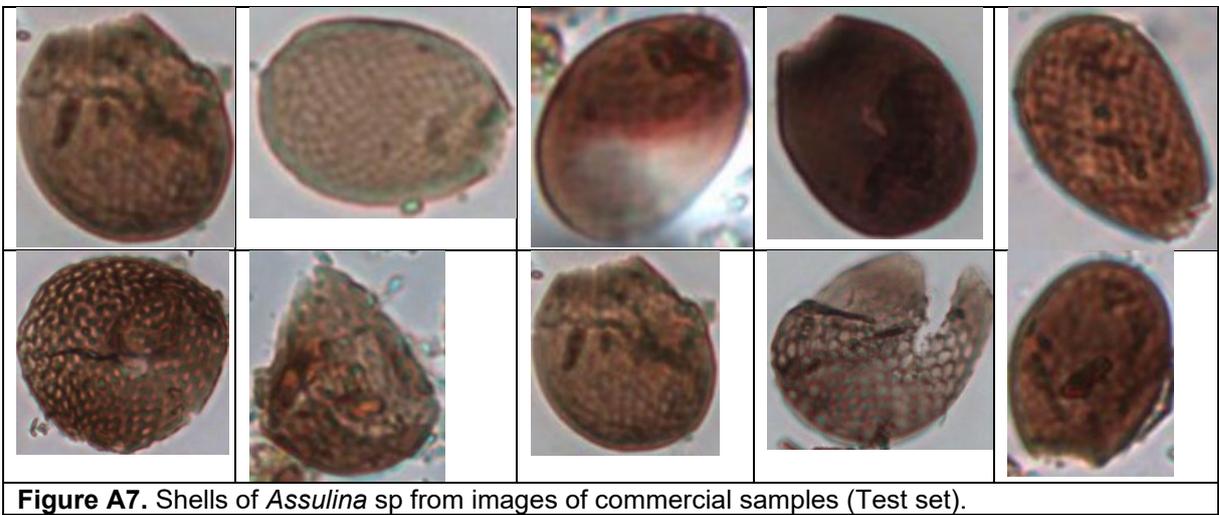
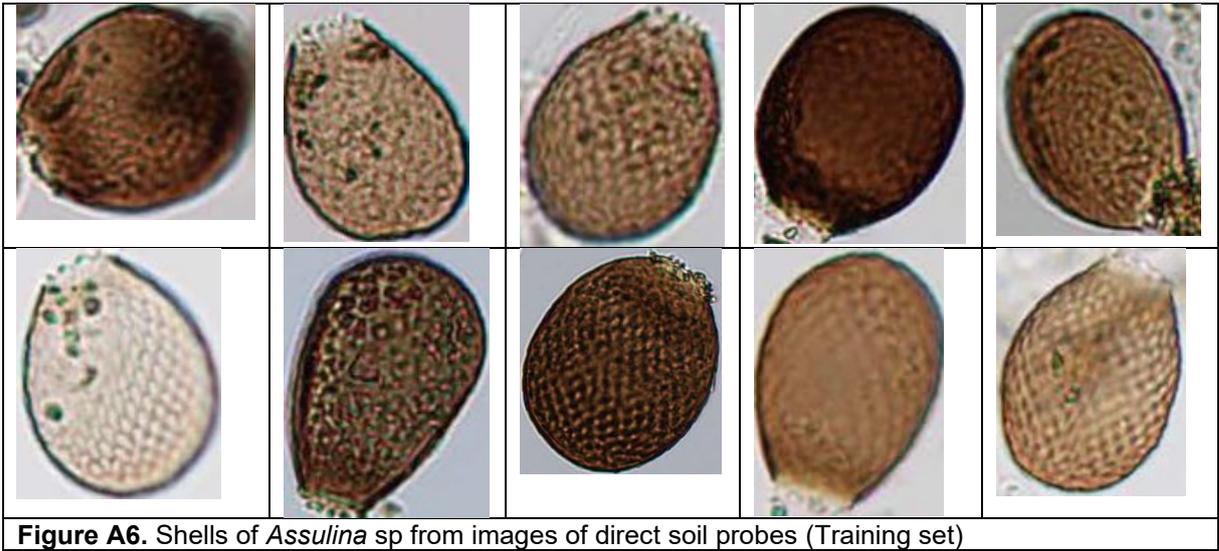


Figure A5. Shells of *Amphitrema* sp from images of commercial samples (Test set).



TRAINING : Direct soil samples		VALIDATION : Commercial samples			TEST : Commercial samples	
Description	Sample-ID	Description	Sample-ID	CV-split	Description	Sample-ID
PEAT 1666 images	4454	PEAT 2554 images grid 1705 active C 283 active R 283 active L 283	5868	fold 0	PEAT 2015 images grid 1200 active C 272 active R 271 active L 272	6241
	4457		5871	fold 0		6242
	4459		6000	fold 0		6243
	4460		5875	fold 1		6244
	4461		6041	fold 1		6251
	4463		6158	fold 1		6252
	4542		5996	fold 2		6253
	4551		5998	fold 2		6254
	4555		5874	fold 2		6267
	4558		5872	fold 3		6268
	4559	5873	fold 3	6273		
	4572	6150	fold 0	6282		
	4573	5997	fold 1	6160		
	4574	5863	fold 2	6164		
	4575	6040	fold 2	6169		
4579	5864	fold 3	6171			
	5999	fold 3				
		NON-PEAT 901 images only grid			NON-PEAT 400 images only grid	

Figure A8. Overview off all samples used in the study and their role. C: Center; R: Right; L: Left

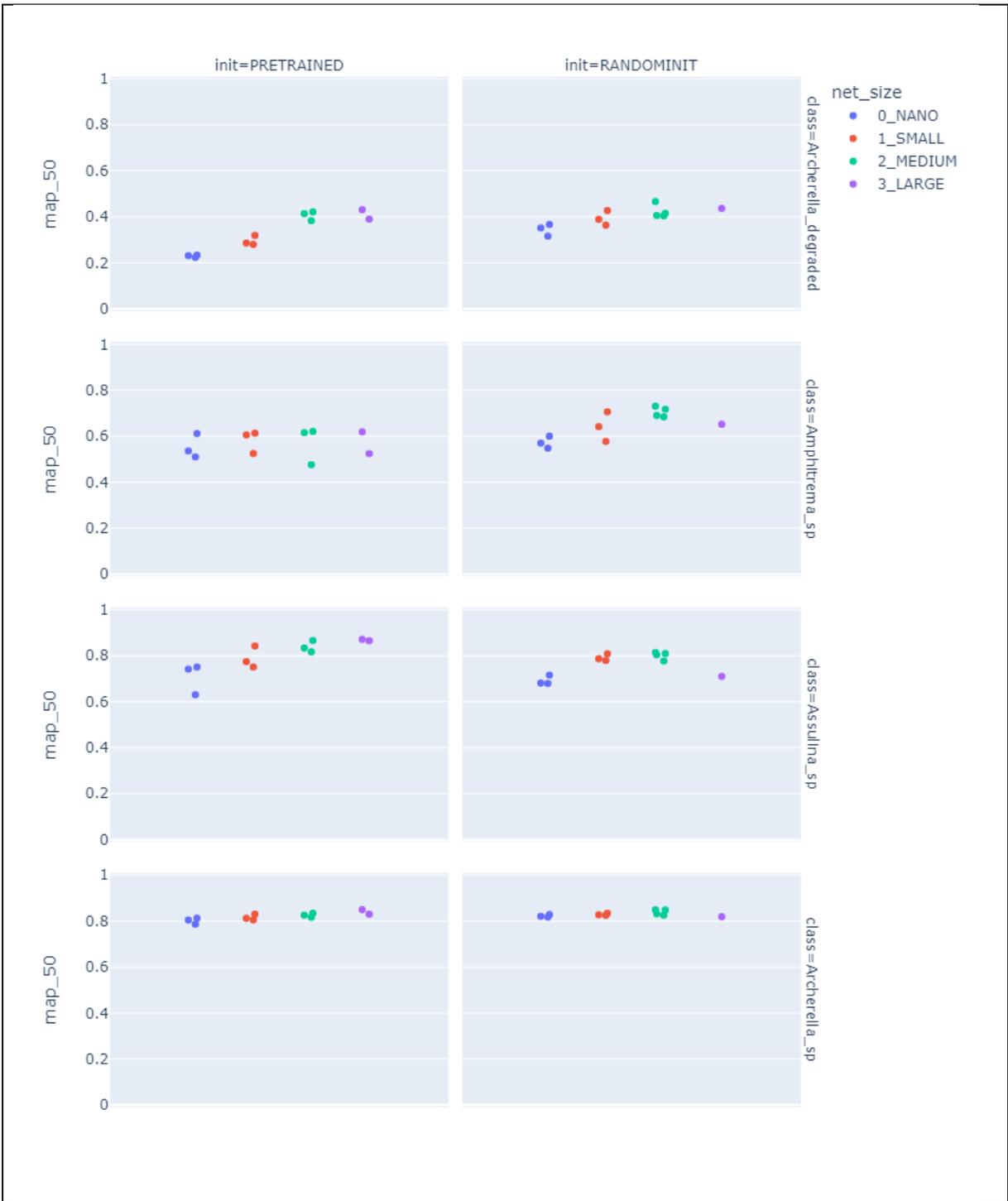


Figure A9. Comparison of detection performance (map_50) for several YOLO model sizes (NANO, SMALL, MEDIUM, LARGE) and weight initialization methods (PRETRAINED, RANDOMINIT).

	Sample-ID	nb_images_total	Amphitrema_sp	Archerella_degraded	Archerella_sp	Assulina_sp	Euglypha_sp	Heleopera_sphagni	Hyalosphenia_elegans_aggr	Hyalosphenia_papilio	Nebelia_combined	Pianocarina_carinata
Training (direct peat samples)	4454	111	0	0	9	2	47	0	10	24	53	8
	4457	111	1	0	15	7	11	0	38	7	39	1
	4459	100	0	0	0	35	0	0	1	0	0	0
	4460	110	0	0	0	7	31	0	3	23	22	0
	4461	113	1	0	2	37	31	0	20	0	7	0
	4463	110	0	0	0	50	31	0	2	0	20	0
	4542	110	1	0	0	11	17	0	0	13	28	0
	4551	74	1	0	3	5	24	1	2	10	9	0
	4555	112	6	0	12	11	12	2	14	35	1	3
	4558	106	36	0	13	4	55	4	0	3	0	6
	4559	110	0	0	28	19	29	8	8	15	12	0
	4572	102	0	0	4	37	18	6	1	9	8	0
	4573	102	20	0	19	0	39	1	0	19	0	8
	4574	101	0	0	21	20	18	13	9	10	1	0
	4575	104	0	0	17	12	17	1	12	20	11	0
	4579	90	14	0	2	6	25	0	1	13	10	1
Total		1666	80	0	145	263	405	36	121	201	221	27

Table A1. Detailed testate amoeba individual counts of the Training

dataset name	Sample-ID	Type	proportion of peat	Type detail	geographical site
Validation set	5868	peat	100%	gardening peat	N.D.
	5871	peat	100%	peat for potting soils	N.D.
	5872	peat	100%	peat for potting soils	N.D.
	5874	peat	100%	peat for potting soils	Latvia
	5875	peat	100%	peat for potting soils	N.D.
	6000	peat	100%	black peat	N.D.
	6041	peat	100%	peat	N.D.
	5873	peat	100%	peat for potting soils	Latvia
	5996	peat	100%	white peat	Estonia
	5998	peat	100%	peat	N.D.
	6158	peat	100%	sphagnum moss for terrarium	N.D.
	5863	non-peat	0%	compost for potting soils	N.D.
	5864	non-peat	0%	wood fibres	N.D.
	5997	non-peat	0%	horticultural growing media	N.D.
	5999	non-peat	0%	universal growing media	N.D.
	6040	non-peat	0%	universal growing media	N.D.
6150	non-peat	0%	universal potting soil	N.D.	
Test set	6241	peat	100%	white peat	N.D.
	6242	peat	100%	black peat	N.D.
	6243	peat	100%	black peat	Estonia
	6244	peat	100%	white peat	Estonia
	6251	peat	100%	white peat	Lithuania
	6252	peat	100%	white peat	Germany
	6253	peat	100%	peat	Lithuania / Latvia
	6254	peat	100%	peat	Lithuania / Latvia
	6267	peat	100%	white peat	Germany / East. Eur.
	6268	peat	100%	white peat	Baltic countries
	6273	peat	100%	black peat	Lithuania
	6282	peat	100%	black peat	Germany
	6160	non-peat	0%	universal potting soil	N.D.
	6164	non-peat	0%	potting soil for seedlings	N.D.
6169	non-peat	0%	potting soil for potted plants	N.D.	
6171	non-peat	0%	garden compost	Switzerland	

Table A2. details on all commercial samples used in this study

		grid search										active search center													
	Sample-ID	nb_images_total	Amphitrema_sp	Archerella_degraded	Archerella_sp	Assulina_sp	Euglypha_sp	Heleopera_sphagni	Hyalosphenia_elegans_aggr	Hyalosphenia_papilio	Nebela_combined	Planocarina_carinata	nb_images_total	Amphitrema_sp	Archerella_degraded	Archerella_sp	Assulina_sp	Euglypha_sp	Heleopera_sphagni	Hyalosphenia_elegans_aggr	Hyalosphenia_papilio	Nebela_combined	Planocarina_carinata		
Validation (non-peat)	5863	150	0	0	0	0	0	0	0	0	0	0	not applicable												
	5864	150	0	0	0	0	0	0	0	0	0	0													
	5997	151	0	0	0	0	0	0	0	0	0	0													
	5999	150	0	0	0	0	0	0	0	0	0	0													
	6150	150	0	0	0	0	0	0	0	0	0	0													
	6040	150	0	0	0	0	0	0	0	0	0	0													
Validation (peat)	5868	153	1	0	1	0	0	0	0	0	0	0	11	0	4	4	0	0	0	0	1	0	0	0	
	5871	200	3	6	14	1	0	0	0	1	0	0	35	4	6	13	4	0	1	0	1	0	0	0	
	5872	150	0	2	4	1	0	0	0	0	0	0	31	0	6	11	2	0	0	0	2	0	0	0	
	5873	150	0	3	2	0	0	0	0	0	0	0	22	0	3	14	2	0	1	0	0	0	0	0	
	5874	150	1	3	4	2	0	0	0	0	0	0	33	1	2	21	2	0	0	0	2	0	0	0	
	5875	150	0	1	7	0	0	0	0	0	0	0	48	3	16	18	4	0	1	1	0	0	0	0	
	5996	150	0	1	2	0	0	0	0	0	0	0	21	1	6	4	1	0	0	0	0	0	0	0	
	5998	150	0	0	2	0	0	0	0	0	0	0	10	0	0	4	1	0	0	0	0	0	0	0	
	6000	149	0	1	5	0	0	0	0	0	0	0	29	5	8	10	3	0	0	0	0	0	0	0	
	6041	150	0	0	4	0	0	0	0	0	0	0	24	1	6	10	0	0	0	0	0	0	0	0	
	6158	153	0	0	0	3	1	0	0	0	0	0	19	0	0	0	3	8	0	0	0	0	0	0	
	Total peat		1705	5	17	45	7	1	0	0	1	0	0	283	15	57	109	22	8	3	1	6	0	0	

Table A3. Detailed testate amoeba individual counts of the Validation set.

		grid search										active search center													
	Sample-ID	nb_images_total	Amphitrema_sp	Archerella_degraded	Archerella_sp	Assulina_sp	Euglypha_sp	Heleopera_sphagni	Hyalosphenia_elegans_aggr	Hyalosphenia_papilio	Nebela_combined	Planocarina_carinata	nb_images_total	Amphitrema_sp	Archerella_degraded	Archerella_sp	Assulina_sp	Euglypha_sp	Heleopera_sphagni	Hyalosphenia_elegans_aggr	Hyalosphenia_papilio	Nebela_combined	Planocarina_carinata		
Test (non-peat)	6160	100	0	0	0	0	0	0	0	0	0	0	not applicable												
	6164	100	0	0	0	0	0	0	0	0	0	0													
	6169	100	0	0	0	0	0	0	0	0	0	0													
	6171	100	0	0	0	0	0	0	0	0	0	0													
Test (peat)	6241	100	1	1	1	0	0	0	0	0	0	0	30	3	3	16	1	0	0	1	0	0	0	0	
	6242	100	0	0	1	0	0	0	0	0	0	0	31	4	6	14	4	0	0	0	0	0	0	0	
	6243	100	0	0	0	0	0	0	0	0	0	0	7	0	1	4	1	0	0	0	0	0	0	0	
	6244	100	0	0	0	0	0	0	0	0	0	0	10	0	0	6	1	0	0	0	0	0	0	0	
	6251	100	0	0	2	0	0	0	0	0	0	0	25	1	5	15	0	0	0	0	0	0	0	0	
	6252	100	0	0	0	0	0	0	0	0	0	0	16	0	1	6	2	0	0	0	0	1	0	0	
	6253	100	2	1	0	0	0	0	0	0	0	0	27	5	7	11	1	1	0	0	0	0	0	0	
	6254	100	0	0	2	0	0	0	0	0	0	0	42	0	11	23	3	0	0	1	2	0	0	0	
	6267	100	1	0	4	0	0	0	0	0	0	0	31	1	4	23	2	0	0	1	0	0	0	0	
	6268	100	0	0	4	0	0	0	0	0	0	0	26	3	4	13	1	0	0	1	3	0	0	0	
	6273	100	0	0	0	0	0	0	0	0	0	0	3	0	0	1	0	0	0	0	0	1	0	0	
	6282	100	1	0	3	0	0	0	0	0	0	0	24	2	1	17	3	0	0	0	0	0	0	0	
Total peat		1200	5	2	17	0	0	0	0	0	0	272	19	43	149	19	1	0	4	7	0	0	0		

Table A4. Detailed testate amoeba individual counts of the Test set.