Pleurotus Mushroom Dataset

Readme File

ITHACA – University of Western Macedonia - https://ithaca.ece.uowm.gr/

Authors: Vasileios Moysiadis, Christos Chaschatzis, Anna Triantafyllou, Athanasios Liatifis, Ilias Siniosoglou, Vicky Kelly, Chrysoula Karaiskou, Dimitrios Tsiamitros, Georgios Kakamoukas, Athina Krestou, Theofano Kollatou, Georgios Kokkonis, Stamatia Bibi, Ioannis Moscholios, Nikolaos Maropoulos, Panagiotis Sarigiannidis

Publication Date: May 19, 2025

This research was cofunded by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Western Macedonia 2014–2020, under the call "Collaborative and networking actions between research institutions, educational institutions and companies in priority areas of the strategic smart specialization plan of the region", project "Smart Mushroom fARming with internet of Things—SMART", project code: DMR-0016521.

1. Abstract

This dataset comprises two independent datasets related to Pleurotus mushrooms, designed to monitor their growth and determine when they are ready for harvest. The data were collected from a greenhouse in Grevena, Greece, during July and August 2022.

The first dataset, “Mushroom Harvest” was annotated under the guidance of an expert agriculturist and categorizes mushroom growth into three distinct classes. This dataset is intended to identify mushrooms that are ready for harvest.

The second dataset, “Mushroom Growth”, consists of a series of photos taken from 33 different substrate grow bags for seven different days. This dataset is suitable for monitoring the growth rate of mushrooms and determining when they are ready for harvest.

2. Citation

Please cite the following papers when using this dataset:

V. Moysiadis, G. Kokkonis, S. Bibi, I. Moscholios, N. Maropoulos, and P. Sarigiannidis, “Monitoring Mushroom Growth with Machine Learning,” Agriculture, vol. 13, no. 1, p. 223, Jan. 2023, doi: 10.3390/agriculture13010223.

V. Moysiadis, P. Sarigiannidis, V. Vitsas, and A. Khelifi, “Smart farming in Europe,” Computer Science Review, vol. 39, p. 100345, Jan. 2021, doi: 10.1016/j.cosrev.2020.100345.

3. Dataset Collection

First Dataset – Mushroom Harvest

The photos in the first dataset were collected using a smartphone in a greenhouse in Grevena, Greece, during various mushroom cultivation cycles in 2022. The annotation was performed under the guidance of an expert agriculturist, and the images were categorized into three classes—Stage1, Stage2, and Stage3—based on the growth stage, as follows:

Stage 1: This stage corresponds to the initial days of appearance when the mushrooms were very small and had just begun to form their shape.

Stage 2: This stage includes mushrooms that had already formed their shape and were continuing to grow. Some were still visibly small, while others were larger but not yet ready for harvest.

Stage 3: This stage corresponds to mushrooms that were ready for harvest. This readiness could be determined by two key indicators. Specifically, when the edges of their caps started to flatten or slightly curl upward, it signified the final stage of growth [1].

Second Dataset – Mushroom Growth

The second dataset was created to monitor the growth rate of mushrooms. The photos were collected using a smartphone in a greenhouse in Grevena, Greece. The distance between the camera and the substrate grow bag was one meter. More specifically, 231 photos were collected from 33 different mushroom substrate grow bags for seven different days, between August 16, 2022, and August 24, 2022. All photos have a resolution of 3000 × 3000 pixels.

4. Annotation

First Dataset – Mushroom Harvest

The annotation was implemented with MakeSense.ai online tool, and exported in YOLO (You Only Look Once) format with one .txt files for each photo. The YOLO format is commonly used for object detection tasks. It represents annotations in a text file, where each object in an image is described by a line in the file. Each line in the annotation file represents one object in the image. The format for each line is:

```
<class_id> <x_center> <y_center> <width> <height>
```

Where:

class_id: The integer index (starting from 0) of the class to which the object belongs.

x_center: The x-coordinate of the center of the bounding box, normalized by the width of the image. This value is a float between 0 and 1.

y_center: The y-coordinate of the center of the bounding box, normalized by the height of the image. This value is a float between 0 and 1.

width: The width of the bounding box, normalized by the width of the image. This value is a float between 0 and 1.

height: The height of the bounding box, normalized by the height of the image. This value is a float between 0 and 1.

Each annotated photo (Fig. 1) may contain one or multiple mushrooms from one or more classes.

Fig. 1. Annotated photos with mushrooms in different classes.

The dataset overview is as follows:

Annotated photos	1128
Annotated mushrooms	4271
Growing Stage1	1130
Growing Stage2	1845
Growing Stage3	1296

Second Dataset – Mushroom Growth

The annotation of the photos is indicated in the filenames and folder names (Fig. 2). For instance, the folder named S07 corresponds to substrate grow bag number 7 and contains all the photos for that specific grow bag. Additionally, the filenames of the photos have the following format:

IMG_YYYYMMDD_HHMMSS.jpg

where:

YYYY: Corresponds to the year (e.g., 2022)

MM: Corresponds to the month (e.g., 08 for August)

DD: Corresponds to the day of the month (e.g., 19)

HH: Corresponds to the hour (e.g., 14)

MM: Corresponds to the minutes

HH: Corresponds to the seconds

IMG_20220818_130747.jpg
Captured on 18/8/2022

IMG_20220820_123752.jpg
Captured on 19/8/2022

IMG_20220822_123150.jpg
Captured on 22/8/2022

Fig. 1. Photos of the same substrate growth bag in three different days.

Note: Not all substrate grow bag numbers are available, as some of them did not produce any mushrooms during this period.

5. Structure and Format

The provided dataset is organized with the following structure:

```
Pleurotus Mushroom Dataset
|   ReadMe.pdf
+---mh-dataset
|   +---images
|   |   |   filename-1.jpg
|   |   |   filename-2.jpg
|   |   |   <..>
|   +---labels
|   |   |   labels.txt
|   |   |   filename-1.txt
|   |   |   filename-2.txt
|   |   |   <..>
+---mg-dataset
|   +---S01
|   |   | IMG_YYYYMMDD_HHMMSS.jpg
|   |   | <...>
|   +---S02
|   |   | IMG_YYYYMMDD_HHMMSS.jpg
|   |   | <...>
|   +---<..>
|   +---<..>
|   +---S37
|   |   | IMG_YYYYMMDD_HHMMSS.jpg
|   |   | <...>
```

Name	Type	Property
ReadMe.pdf	Report	A file that contains the documentation of the dataset
mh-dataset	Directory	A folder containing the dataset for mushroom harvest
images	Directory	A folder containing the images of the dataset for mushroom harvest
filename.jpg	Image File	A photo with mushrooms in the greenhouse
labels	Directory	A folder containing the annotation files of the dataset for mushroom harvest
filename.txt	Text File	The annotation for the photo “filename.jpg”
labels.txt	Text File	A text file that contains the labels of the classes
mg-dataset	Directory	A folder containing the dataset for mushroom growth
S01	Directory	A folder containing the images of the annotated files for substrate grow bag number 1.
`IMG_YYYYMMDD_HHMMSS.jpg`	Image File	A photo of a substrate grow bag captured on specific day and time.

6. Acknowledgment

This research was cofunded by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Western Macedonia 2014–2020, under the call "Collaborative and networking actions between research institutions, educational institutions and companies in priority areas of the strategic smart specialization plan of the region", project "Smart Mushroom fARming with internet of Things—SMART", project code: DMR-0016521.

7. References

[1] W. Yang, F. Guo, and Z. Wan, "Yield and size of oyster mushroom grown on rice/wheat straw basal substrate supplemented with cotton seed hull," Saudi Journal of Biological Sciences, vol. 20, no. 4, pp. 333–338, Mar. 2013, doi: 10.1016/j.sjbs.2013.02.006.