# Cherry Tree Disease Detection Dataset

## Readme File

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

**Authors:** Christos Chaschatzis, Ilias Siniosoglou, Anna Triantafyllou, Chrysoula Karaiskou, Athanasios Liatifis, Panagiotis Radoglou-Grammatikis, Dimitrios Pliatsios, Vasiliki Kelli, Thomas Lagkas, Vasileios Argyriou, Panagiotis Sarigiannidis

**Publication Date:** September 26, 2022

This work was co-financed by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (project code: T1EDK-04759).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No. 957406 (TERMINET).

# 1. Introduction

This cherry tree disease detection dataset is a multimodal, multi-angle dataset which was constructed for monitoring the growth of cherry trees, including stress analysis and prediction. An orchard of cherry trees is considered in the area of Western Macedonia, where 577 cherry trees were recorded in a full crop season starting from Jul. 2021 to Jul. 2022. The dataset includes a) aerial / Unmanned Aerial Vehicle (UAV) images, b) ground RGB images/photos, and c) ground multispectral images/photos. Two agronomist experts annotated the dataset by identifying a stress, which in this case is a common disease in cherry trees known as Armillaria [1][2].

## 2. Citation

Please cite the following papers when using this dataset:

C. Chaschatzis, C. Karaiskou, E. Mouratidis, E. Karagiannis, and P. Sarigiannidis, "Detection and Characterization of Stressed Sweet Cherry Tissues Using Machine Learning", Drones, vol. 6, no. 1, 2022.

P. Radoglou-Grammatikis, P. Sarigiannidis, T. Lagkas, & I. Moscholios, "A compilation of UAV applications for precision agriculture," Computer Networks, vol. 172, no. 107148, 2020.

A. Lytos, T. Lagkas, P. Sarigiannidis, M. Zervakis, & G. Livanos, "Towards smart farming: Systems, frameworks and exploitation of multiple sources," Computer Networks, vol. 172, no. 107147, 2020.

# 3. Cherry tree mapping

In this dataset, an orchard of cherry trees is considered in the area of Western Macedonia, where 577 cherry trees were recorded in a full crop season starting from Jul. 2021 to Jul. 2022. The tree mapping within the orchard is depicted in Fig. 1., where each circle represents a cherry tree. Labels on the circles (green, red etc) will be elaborated in the following Sections. The five time periods, where the orchard was recorded are: 8th of Jul. 2021, 16th of Sep. 2021, 3rd of Nov. 2021, 26th of May 2022, and 13th of Jul. 2022, providing data to a full year of life cycle.

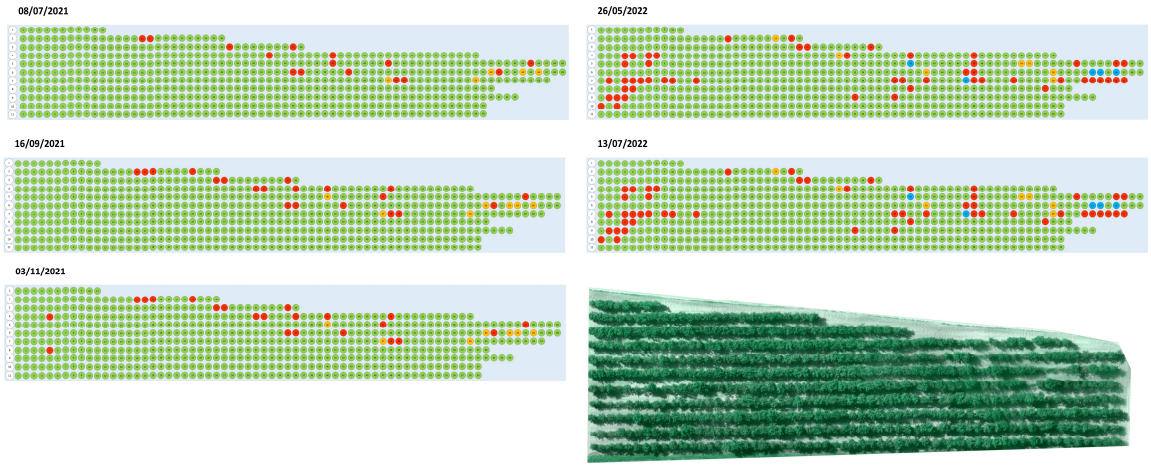



**Fig. 1. Orchard mapping using annotated circles, one for each cherry tree, in five total crop periods.**

# 4. Dataset Modalities

The dataset includes a) aerial / Unmanned Aerial Vehicle (UAV) images, b) ground RGB images/photos, and c) ground multispectral images/photos. Two agronomist experts annotated the dataset by identifying a stress, which in this case is a common disease in cherry trees known as Armillaria [1][2]. In particular, the following modalities are featured in the dataset:

- Ground RGB images
- Ground multispectral images
- UAV/Aerial images (RGB, multispectral, and NDVI).

These modalities represent the cherry tree cultivation in many levels. Each modality describes the same object (cherry tree) within the dataset, i.e., for each tree within. For example, Fig. 2 show RGB images, Fig. 3 illustrates multispectral images, and Fig. 4 provides UAV images. All images show the same cherry trees under three (RGB, multispectral, and UAV) aspects.

**Fig. 2. RGB Images**

**Fig. 3. Multispectral images**

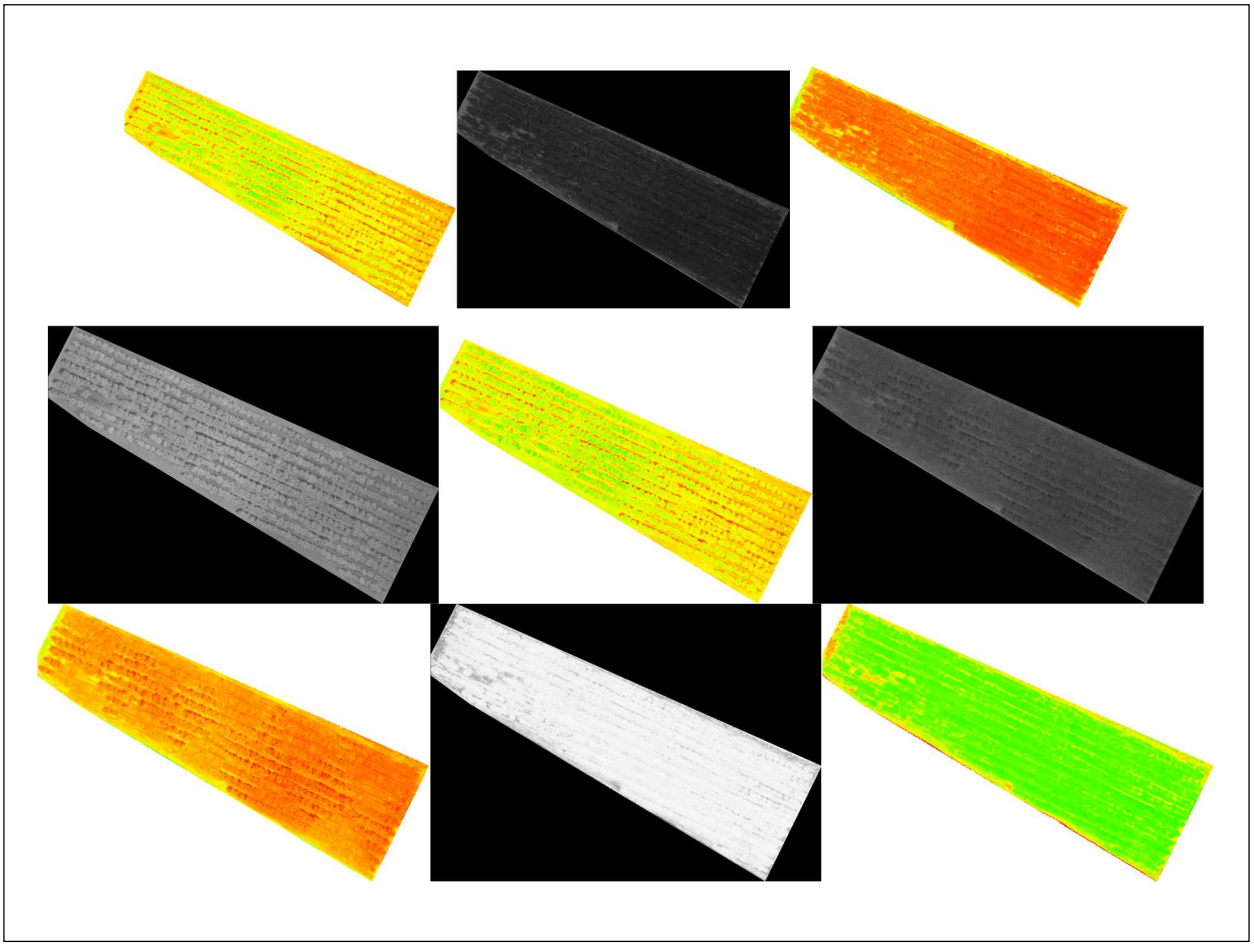


**Fig. 4. UAV/Aerial Images (multispectral/NDVI)**

# 5. Dataset Collection & Annotation

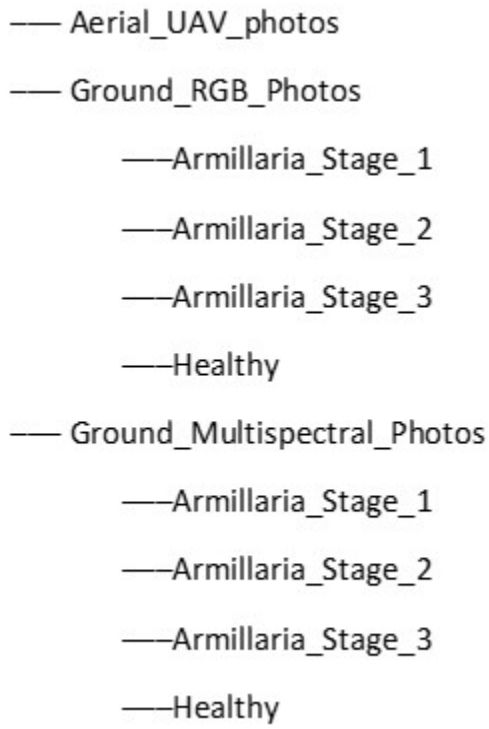
This dataset was annotated by two agronomist experts in terms of disease stage (Armillaria). In particular, they annotated each cherry tree, one by one, in four levels of disease stage:

- Healthy: the cherry tree is completely healthy;
- Stage1: Armillaria is present in light form in the cherry tree;
- Stage2: Armillaria is present in advanced form;
- Stage3: the cherry tree is killed due to Armillaria.

The annotation process was considered by each one of the underlying modalities (RGB, multispectral and UAV/aerial).

## 5.1 Image Collection

The image collection is depicted in the following image in terms of the three modalities (aerial / Unmanned Aerial Vehicle (UAV) images, ground RGB images/photos, and ground multispectral images/photos).



## 5.2 Dataset Overview

The dataset overview is depicted in Table 1.

| Date | Status | Number of RGB images (once per tree) | Number of multispectral images (five images per tree) |
|---|---|---|---|
| **08/07/2021** | Armillaria Stage 1 | 15 | 75 |
| | Armillaria Stage 2 | 7 | 35 |
| | Armillaria Stage 3 | 0 | 0 |
| | Healthy | 555 | 2775 |
| **16/09/2021** | Armillaria Stage 1 | 22 | 110 |
| | Armillaria Stage 2 | 7 | 35 |
| | Armillaria Stage 3 | 0 | 0 |
| | Healthy | 548 | 2730 |
| **03/11/2021** | Armillaria Stage 1 | 22 | 110 |
| | Armillaria Stage 2 | 7 | 35 |
| | Armillaria Stage 3 | 0 | 0 |
| | Healthy | 548 | 2730 |
| **26/05/2022** | Armillaria Stage 1 | 52 | 255 |
| | Armillaria Stage 2 | 7 | 35 |
| | Armillaria Stage 3 | 5 | 0 |
| | Healthy | 510 | 2555 |
| **13/07/2021** | Armillaria Stage 1 | 51 | 255 |
| | Armillaria Stage 2 | 7 | 35 |
| | Armillaria Stage 3 | 5 | 25 |
| | Healthy | 514 | 2852 |

**Table 1. Dataset overview**

# 6. Structure and Format

## 6.1 Dataset Structure

The provided dataset has the following structure:

```
Cherry Tree Disease Detection Dataset
|   Readme.docx
|   Tree_Mapping.xls
|
+---Date
|   |   date.csv
|   |
|   +---Aerial_UAV_photos
|   |   |       Image01
|   |   |       Image02
|   |   |         <…>
|   |   |
|   +---Ground_RGB_Photos
|   |   +---Armillaria_Stage_1
|   |   |       Image01
|   |   |       Image02
|   |   |         <…>
|   |   +---Armillaria_Stage_2
|   |   |       Image01
|   |   |       Image02
|   |   |         <…>
|   |   +---Armillaria_Stage_3
|   |   |       Image01
|   |   |       Image02
|   |   |         <…>
|   |   \---Healthy
|   |   |       Image01
|   |   |       Image02
|   |   |         <…>
|   |   |
|   \---Ground_Multispectral_Photos
|       +---Armillaria_Stage_1
|       |   +---Folder01
|       |   |       Image01
|       |   |       Image01
|       |   |       MetadataFile01
|       |   |         <…>
|
```

Where:

| Name | Type | Property |
|---|---|---|
| Readme.docx | Report | A File that contains the documentation of the ERICA Dataset |
| Tree_Mapping.xlsx | Report | An Excel file that contains the mapping of Date of collection, Trees |

| | | |
|---|---|---|
| | | (row/column) in a cultivation and colouring of the trees depending on their Armillaria infection stage. Shortly:<br>● -> Healthy<br>● -> Armillaria Stage 1<br>● -> Armillaria Stage 2<br>● -> Armillaria Stage 3 |
| Date | Directory | A folder that contains the data collected on a certain date |
| date.csv | Annotation File | A file that contains the label of the Armillaria infection stage per Cherry Tree for a specific date of collection. Shortly the labels are:<br><br>0 -> Healthy<br>1 -> Armillaria Stage 1<br>2 -> Armillaria Stage 2<br>3 -> Armillaria Stage 3 |
| Aerial_photos* | Directory | A folder that contains images of Aerial photos from the Cherry Tree cultivation. |
| Ground_Photos | Directory | A folder that contains RGB images of each tree separately from the Cherry Tree cultivation. The images are named as <Row-Column>.ext |
| Multispectral | Directory | A folder that contains folders of Multispectral images of each tree separately from the Cherry Tree cultivation. The folder contained in Multispectral is named as <Row-Column> |
| Armillaria_Stage_1 | Directory | A folder containing images of Cherry Trees within the Armillaria Stage 1 |
| Armillaria_Stage_2 | Directory | A folder containing images of Cherry Trees within the Armillaria Stage 2 |
| Armillaria_Stage_3 | Directory | A folder containing images of Cherry |

| | | |
|---|---|---|
| | | Trees within the Armillaria Stage 3 |
| Healthy | Directory | A folder containing images of Cherry Trees that are Healthy |

*The provided Aerial Images have the .tif format. To open these images, follow the guide below.

### 6.2 Guide to edit the *.tif files

The Aerial/UAV images contain images obtained from the UAV camera in the .tif format. To open these images, you will need the QGIS or other relevant program, or load them by using the corresponding python libraries. Please follow the steps below:

1. Open QGIS
2. Locate the browser window in QGIS
3. Navigate to the folder that contains the images and select all the images in the layer.
4. Once you have selected the images, select Add Layer to Project, and the selected image will be added to your map.

For accessing the Image data with the OpenCV python library the following code example is provided:

```
import sys
import math
import cv2 as cv
import numpy as np
from PIL import Image, ImageOps, ImageChops
import PIL
import matplotlib.pyplot as plt

red = Image.open("arnisa_red.tif")

nir = Image.open("arnisa_nir.tif")

rededge = Image.open("arnissa_rededge.tif")

nir = np.asarray(nir)
red = np.asarray(red)
rededge = np.asarray(rededge)

ndvi = np.zeros([nir.shape[0], nir.shape[1], 3], dtype=np.uint8)
ndviData = np.zeros(nir.shape[0])
ndviData = (nir - red) / (nir + red)

for i in range(0, ndviData.shape[0]):
        for j in range(0, ndviData.shape[1]):
        if (ndviData[i][j] < 0):
           ndvi[i][j][0] = 0
           ndvi[i][j][1] = 0
           ndvi[i][j][2] = 0
```

```
    elif (ndviData[i][j] > 0):
        if(ndviData[i][j] > 0 and ndviData[i][j] < 0.06):
              ndvi[i][j][0] =  255
              ndvi[i][j][1] = 0
              ndvi[i][j][2] = 0
           elif (ndviData[i][j] > 0.06 and ndviData[i][j] < 0.12):
              ndvi[i][j][0] =  255
              ndvi[i][j][1] = 32
           ndvi[i][j][2] = 0
           elif (ndviData[i][j] > 0.12 and ndviData[i][j] < 0.18):
              ndvi[i][j][0] =  255
              ndvi[i][j][1] = 64
              ndvi[i][j][2] = 0
           elif (ndviData[i][j] > 0.18 and ndviData[i][j] < 0.24):
              ndvi[i][j][0] =  255
              ndvi[i][j][1] = 96
              ndvi[i][j][2] = 0
           elif (ndviData[i][j] > 0.24 and ndviData[i][j] < 0.30):
              ndvi[i][j][0] =  255
              ndvi[i][j][1] = 128
              ndvi[i][j][2] = 0
           elif (ndviData[i][j] > 0.30 and ndviData[i][j] < 0.36):
              ndvi[i][j][0] =  255
              ndvi[i][j][1] = 160
              ndvi[i][j][2] = 0
           elif (ndviData[i][j] > 0.36 and ndviData[i][j] < 0.42):
              ndvi[i][j][0] =  255
              ndvi[i][j][1] = 192
              ndvi[i][j][2] = 0
           elif (ndviData[i][j] > 0.42 and ndviData[i][j] < 0.48):
               ndvi[i][j][0] =  255
              ndvi[i][j][1] = 224
              ndvi[i][j][2] = 0
           elif (ndviData[i][j] > 0.48 and ndviData[i][j] < 0.54):
              ndvi[i][j][0] =  255
              ndvi[i][j][1] = 240
                ndvi[i][j][2] = 0
           elif (ndviData[i][j] > 0.54 and ndviData[i][j] < 0.60):
              ndvi[i][j][0] =  224
              ndvi[i][j][1] = 255
              ndvi[i][j][2] = 0
           elif (ndviData[i][j] > 0.60 and ndviData[i][j] < 0.66):
              ndvi[i][j][0] =  192
              ndvi[i][j][1] = 255
              ndvi[i][j][2] = 0
           elif (ndviData[i][j] > 0.66 and ndviData[i][j] < 0.72):
              ndvi[i][j][0] =  176
                 ndvi[i][j][1] = 255
              ndvi[i][j][2] = 0
           elif (ndviData[i][j] > 0.72 and ndviData[i][j] < 0.78):
              ndvi[i][j][0] =  160
              ndvi[i][j][1] = 255
              ndvi[i][j][2] = 0
           elif (ndviData[i][j] > 0.78 and ndviData[i][j] < 0.84):
              ndvi[i][j][0] =  144
              ndvi[i][j][1] = 255
              ndvi[i][j][2] = 0
           elif (ndviData[i][j] > 0.84 and ndviData[i][j] < 0.90):
```

```
                    ndvi[i][j][0] =  64
                    ndvi[i][j][1] = 255
                    ndvi[i][j][2] = 0
                elif (ndviData[i][j]> 0.90):
                    ndvi[i][j][0] =  16
                    ndvi[i][j][1] = 255
                    ndvi[i][j][2] = 0


pil_image=Image.fromarray(ndvi)
img = pil_image
img = img.convert("RGBA")
datas = img.getdata()

newData = []
for item in datas:
        if item[0] == 255 and item[1] == 255 and item[2] == 255:
         newData.append((255, 255, 255, 0))
        else:
         newData.append(item)

img.putdata(newData)
img.save("ndvi_agr.png", "PNG")
```

# 7. Acknowledgment

This work was co-financed by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (project code: T1EDK-04759).


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No. 957406 (TERMINET).

## References

[1] Devkota, P.; Iezzoni, A.; Gasic, K.; Reighard, G.; Hammerschmidt, R. Evaluation of the susceptibility of Prunus rootstock genotypes to Armillaria and Desarmillaria species. Eur. J. Plant Pathol. 2020, 158, 177–193.

[2] Devkota, P.; Hammerschmidt, R. "The infection process of Armillaria mellea and Armillaria solidipes". Physiol. Mol. Plant Pathol. 2020, 112, 101543.