



Peach Tree Disease Detection Dataset

Readme File

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

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

Publication Date: November 23, 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 peach tree disease detection dataset is a multimodal, multi-angle dataset which was constructed for monitoring the growth of peach trees, including stress analysis and prediction. An orchard of peach trees is considered in the area of Thessaly, where 889 peach trees were recorded in a full crop season starting from Jul. 2021 to Sep. 2022. The dataset includes a) aerial / Unmanned Aerial Vehicle (UAV) images, b) ground RGB images/photos, and c) ground multispectral images/photos. Two experts (agronomists) annotated the dataset by identifying whether the trees are healthy or stressed, where the identified stress is either *Anarsia lineatella* or *Grapholita molesta*.

2. Citation

Please cite the following papers when using this dataset:

C. Chaschatzis, C. Karaskou, 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. Peach tree mapping

In this dataset, an orchard of peach trees is considered in the area of Thessaly, where 889 peach trees were recorded in a full crop season starting from Jul. 2021 to Sep. 2022. The tree mapping within the orchard is depicted in Fig. 1., where each circle represents a peach tree. The color annotation on the circles (green, red, blue, yellow) will be elaborated in the following Sections. The six time periods, where the orchard was recorded are: 27th of Jul. 2021, 14th of Sep. 2021, 4th of Nov. 2021, 25th of May 2022, 15th of Jul. 2022, and 14th of Sep. 2022, providing data to a full year of trees life cycle. Each tree is identified by two numbers, the row and the column as the tree is mapped in the orchard. For example, the tree (3-26) is located in the 3rd row, 26th column of the orchard, as depicted in the following mapping images. Within the dataset files, each image of this tree (3-26) is identified as 3-26.jpg, subject to the corresponding folder. For example, if this file, 3-26.jpg, is under the folder Ground_RGB_Photos, the photo 3-26.jpg is referred to the RGB ground image of this tree.

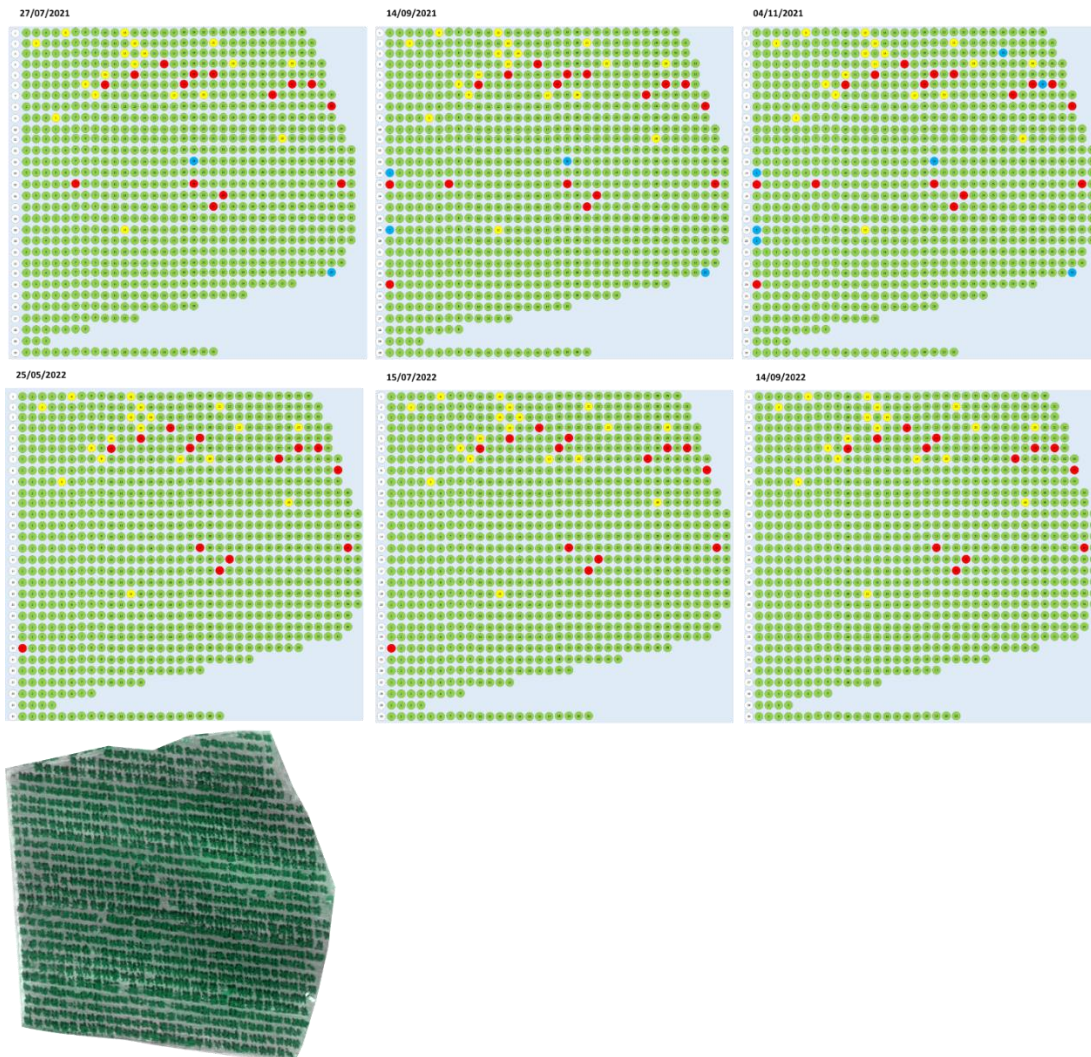


Fig. 1. Orchard mapping using annotated circles, one for each peach tree, covering six periods of time.

4. Dataset Modalities

This dataset includes a) aerial / UAV images, b) ground RGB images/photos, and c) ground multispectral images/photos. Two experts (agronomists) annotated this dataset by identifying stress, which in this case are two common diseases in peach trees: *Grapholita molesta* [1][2] and *Anarsia lineatella* [3]. In particular, the following modalities are featured in this dataset:

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

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



Fig. 2. RGB images



Fig. 3. Multispectral images

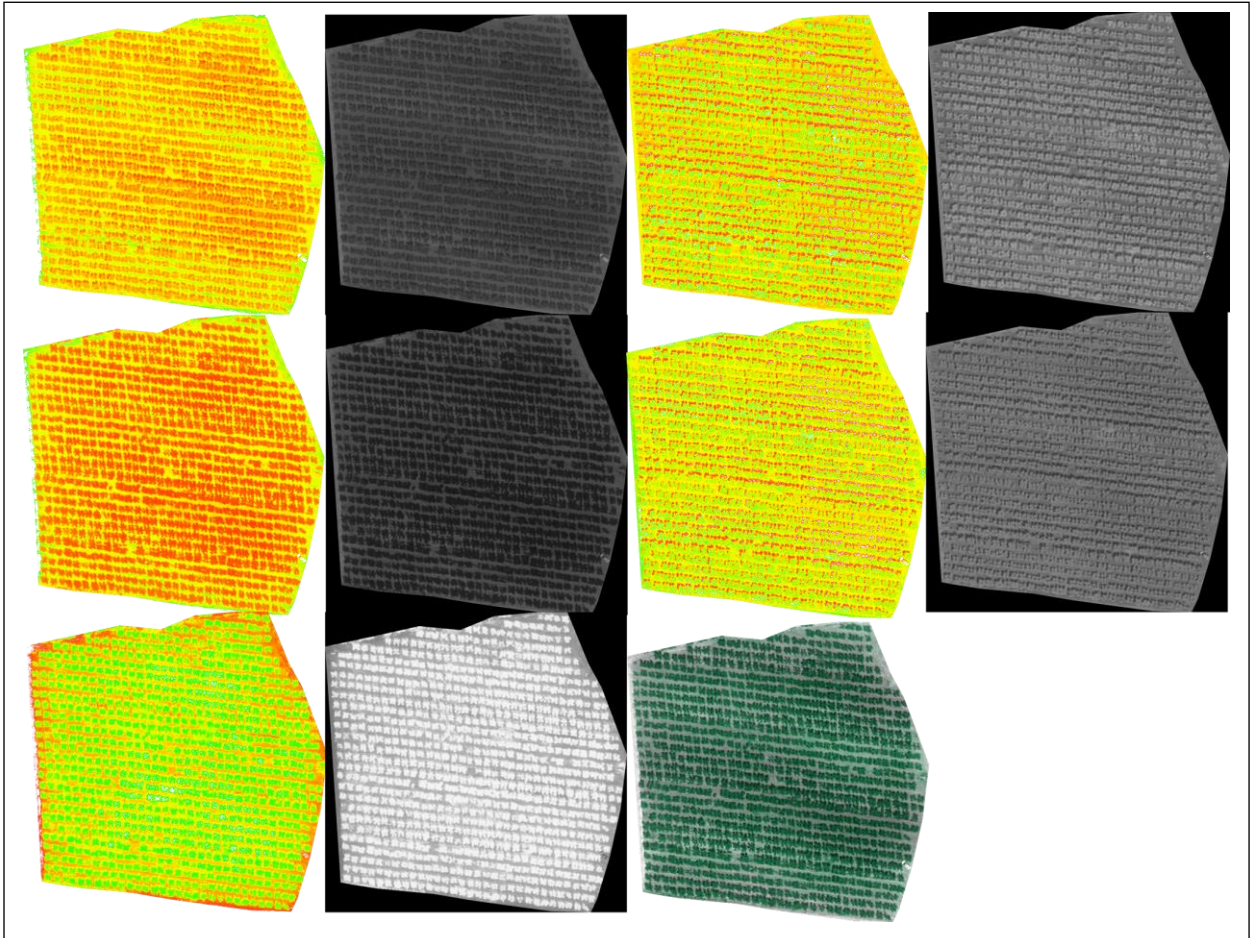


Fig. 4. UAV/Aerial Images (multispectral/NDVI)

5. Dataset Collection & Annotation

This dataset was annotated by two experts (agronomists) in terms of two diseases (Anarsia lineatella and Grapholita molesta). In particular, they annotated each peach tree, one by one, in four states:

- Healthy: the peach tree is completely healthy;
- Anarsia lineatella: the Anarsia lineatella is presented on withered tops on the branches of trees;
- Grapholita molesta: the Grapholita molesta located in the tree trunks and branches of trees;
- Dead Trees: the peach tree is killed due to Anarsia lineatella or Grapholita molesta.

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 figure in terms of the three modalities (aerial / Unmanned Aerial Vehicle (UAV) images, ground RGB images/photos, and ground multispectral images/photos).

--- Aerial_UAV_Photos

--- Ground_RGB_Photos

--- Grapholita molesta

--- Anarsia lineatella

--- Dead trees

--- Healthy

--- Ground_Multispectral_Photos

--- Grapholita molesta

--- Anarsia lineatella

--- Dead trees

--- Healthy

5.2 Dataset Overview

The dataset overview is depicted in Table 1.

Table 1. Dataset overview

Date	Status	Number of RGB images (once per tree)	Number of multispectral images (five images per tree)
27/07/2021	Anarsia lineatella	18	90
	Grapholita molesta	17	85
	Dead Trees	4	20
	Healthy	850	4250
14/09/2021	Anarsia lineatella	18	90
	Grapholita molesta	17	85
	Dead Trees	4	20
	Healthy	850	4250
04/11/2021	Anarsia lineatella	18	90
	Grapholita molesta	17	85
	Dead Trees	7	35
	Healthy	847	4235
25/05/2022	Anarsia lineatella	18	90
	Grapholita molesta	14	70
	Dead Trees	0	0
	Healthy	857	4285
15/07/2022	Anarsia lineatella	18	90
	Grapholita molesta	14	70
	Dead Trees	0	0
	Healthy	857	4285
14/09/2022	Anarsia lineatella	18	90
	Grapholita molesta	13	65
	Dead Trees	0	0
	Healthy	858	4290

Please note that some trees were replaced in the middle of the monitoring period (25/5/2022 and 04/10/2022) since these trees were wither killed or heavily infected by *Anarsia lineatella* or *Grapholita molesta*. The trees that have been replaced are depicted in Table 2. Obviously, these trees were replaced by healthy ones.

Table 2: Trees replaced

Date	Mapped tree
25/05/2022	3-26
	5-21
	6-10
	6-30
	13-19
	14-1
	15-1
	15-7
	19-1
	20-1
	23-33
04/10/2022	24-1

6. Structure and Format

6.1 Dataset Structure

The provided dataset has the following structure:

```
Peach Tree Disease Detection Dataset
|   Readme.docx
|   Tree_Mapping.xls
|
+---Date
|   |   date.csv
|   |
|   +---Aerial_UAV_Photos
|   |   |   Image01
|   |   |   Image02
|   |   |   <...>
|   |
|   +---Ground_RGB_Photos
|   |   +---Anarsia lineatella
|   |   |   Image01
|   |   |   Image02
|   |   |   <...>
|   |   +---Grapholita molesta
|   |   |   Image01
|   |   |   Image02
|   |   |   <...>
|   |   +---Dead Trees
|   |   |   Image01
|   |   |   Image02
|   |   |   <...>
|   |   \---Healthy
|   |   |   Image01
|   |   |   Image02
|   |   |   <...>
|   |
|   \---Ground_Multispectral_Photos
|       +--- Anarsia lineatella
|           +---Folder01
|           |   Image01
|           |   Image01
|           |   MetadataFile01
|           |   <...>
```

Where:

Name	Type	Property
Readme.docx	Report	A file that contains the documentation of this dataset
Tree_Mapping.xlsx	Report	An Excel file containing the tree mapping in terms of (row-column), coloured based on their status: ● → Healthy ● → Grapholita molesta ● → Anarsia lineatella ● → Dead Trees
Date	Directory	A folder indicating the collection date.
date.csv	Annotation File	Similar to Tree_Mapping one, including labels instead of colors in a csv format: 0 -> Healthy 1 -> Grapholita molesta 2 -> Anarsia lineatella 3 -> Dead Trees
Aerial_photos*	Directory	A folder containing tree images of aerial photos.
Ground_Photos	Directory	A folder containing RGB images of each tree in terms of <Row-Column>.jpg
Multispectral	Directory	A folder containing folders of multispectral images of each tree.
Grapholita molesta	Directory	A folder containing images of infected peach trees by Grapholita molesta
Anarsia lineatella	Directory	A folder containing images of infected peach trees by Anarsia lineatella.
Dead Trees	Directory	A folder containing images of Peach Trees within the Dead Trees.
Healthy	Directory	A folder containing images of healthy peach trees.

*The provided aerial Images having .tif format. To open these images, follow the guide below (6.2).

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("red.tif")
nir = Image.open("nir.tif")
rededge = Image.open("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.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] W. Dalmorra de Souza, T. Bystronski Remboski, M. Sanchotene de Aguiar, and P. R. Ferreira Júnior, ‘A Model for Pest Infestation Prediction in Crops Based on Local Meteorological Monitoring Stations’, in 2017 Sixteenth Mexican International Conference on Artificial Intelligence (MICAI), Oct. 2017, pp. 39–45. doi: 10.1109/MICAI-2017.2017.00015.
- [2] V. A. M. Martins, L. C. Freitas, M. S. de Aguiar, L. B. de Brisolara, and P. R. Ferreira, ‘Deep Learning applied to the Identification of Fruit Fly in Intelligent Traps’, in 2019 IX Brazilian Symposium on Computing Systems Engineering (SBESC), Nov. 2019, pp. 1–8. doi: 10.1109/SBESC49506.2019.9046088.
- [3] P. Damos et al., ‘Degree-day risk thresholds for predicting the occurrence of *Anarsia lineatella*, *Grapholita molesta* and *Adoxophyes orana* in northern Greece peach orchards’, *Plant Protect. Sci.*, vol. 58, no. No. 3, May 2022, pp. 234–244 doi: 10.17221/137/2021-PPS.