

BEN- GURION UNIVERSITY OF THE NEGEV
FACULTY OF ENGINEERING SCIENCES
DEPARTMENT OF INDUSTRIAL ENGINEERING AND
MANAGEMENT

Early detection of grapevine downy mildew using thermal imaging

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE M.Sc. DEGREE

By: Bar Cohen

September 2021

BEN- GURION UNIVERSITY OF THE NEGEV
FACULTY OF ENGINEERING SCIENCES
DEPARTMENT OF INDUSTRIAL ENGINEERING AND
MANAGEMENT

Early detection of grapevine downy mildew using thermal imaging

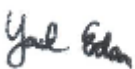
THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE M.Sc. DEGREE

By: Bar Cohen

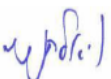
Supervised by: Prof. Yael Edan and Prof. Victor Alchanatis

Author:  _____

Date: 2.10.2021

Supervisor:  _____

Date: 1.10.2021

Supervisor:  _____

Date: 27.9.2021

Chairperson of the Graduate Studies Committee: _____

Date: _____

September 2021

Abstract

Agricultural production must double by 2050 to meet the world's population food demand. Plant diseases are a major threat to global food security and are a major cause of production losses and economic losses in the agriculture industry. Early information on disease development can improve disease control using suitable management strategies.

This research aimed to detect downy mildew on grapevine's leaves at early stages of development by using thermal imaging technology and to determine the best time along the day for image acquisition.

In a controlled experiment conducted in 2019 and 2020, 1403 thermal images of grapevines grown in an experimental greenhouse were acquired before infection and 1, 2, 4, 5, 6, and 7 days after a controlled infection, around midday. In 2020, a second controlled experiment was conducted and images were acquired six days after infection. 575 images of healthy and infected leaves were acquired at seven different times along a single day between 7 a.m. and 4:30 p.m.

Leaves were segmented using the active contour algorithm. Thirteen features were derived from the leaf mask generated by the algorithm and meteorological measurements. Stepwise logistic regression analysis revealed five significant features that were used in five classification models (MTD, STD, perc90, Cv, and CWSI2). The models developed to classify between infected and healthy leaves included decision tree, logistic regression, naive Bayes, support vector machine (SVM), and ensemble. Performance was evaluated using K-folds cross-validation. A sensitivity analysis was conducted for different distributions of data to explain the different results between the days after infection.

The best results were achieved by the SVM model with a classification accuracy of 81.6%, F1 score of 77.5% and AUC of 0.874. Acquiring images in the morning between 10:40 and 11:40 a.m. resulted in 80.7% accuracy, 80.5% F1 score, and 0.895 AUC.

Keywords: Precision agriculture, Disease detection, Classification

Publications

Journal paper

Cohen, B., Edan, Y., Levi, A., & Alchanatis, V. Early detection of grapevine downy mildew and diurnal variations using thermal imaging ([Appendix A](#)).

Submitted October 2021 to Precision agriculture Journal (Q1)

Conference paper and presentation

Cohen, B., Edan, Y., Levi, A., & Alchanatis, V. (2021). Early detection of grapevine downy mildew using thermal imaging. In *Precision agriculture'21* (pp. 283-290). Wageningen Academic Publishers. <https://doi.org/10.3920/978-90-8686-916-9> ([Appendix B](#))

Acknowledgements

Prof. Yael Edan and Prof. Victor Alchanatis, my advisors, were very helpful and this thesis would not have been possible without their insight, support, guidance, and assistance. Financial support for this research was provided by the Israel Innovation Authority, project no 67268.

Partial support was provided by Ben-Gurion University of the Negev through the Rabbi W. Gunther Plaut Chair in Manufacturing Engineering and by the Molcho Center fund.

I am grateful for the support provided by Mr. A. Levi and Mr. S. Ozer during our image acquisition.

I am also grateful for the excellent support of Z. Halamish, N. Guedj, A. Himmel, and A. Movchan from Evogene for growing the plants and assisting with the experiments.

Table of Contents

1. Introduction.....	9
1.1. Problem description.....	9
1.2. Objectives.....	10
1.3. Thesis structure.....	10
2. Literature Review.....	11
2.1. Precision agriculture.....	11
2.2. Thermal sensing.....	12
2.3. Thermal sensing in agriculture.....	13
2.4. Disease detection.....	30
2.5. Vegetational thermal behavior.....	31
2.6. Downy mildew.....	32
2.7. Image capture conditions.....	34
3. Methods.....	36
3.1. Overview.....	36
3.2. Experimental design.....	37
3.2.1. Experiment A.....	37
3.2.2. Experiment B.....	39
3.3. Equipment.....	39
3.4. Databases.....	39
3.4.1. Classification of healthy and infected leaves.....	39
3.4.2. Best acquisition time.....	40
3.5. Segmentation algorithm.....	41
3.6. Features.....	42
3.7. Analysis.....	43
3.8. Classification models.....	44
3.9. Performance measures.....	45
4. Algorithms.....	47
4.1. Edge detection.....	47
4.2. Classification models.....	48
5. Results and Discussion.....	49
5.1. Classification of healthy and infected leaves.....	49
5.1.1. Histograms and boxplots.....	49
5.1.2. Outliers.....	52
5.1.3. Correlation analysis.....	53
5.1.4. Model analysis.....	55
5.1.5. Misclassifications.....	63
5.1.6. Summary.....	64
5.2. Best acquisition time.....	65
5.2.1. Histograms.....	65
5.2.2. Model analysis.....	65
6. Conclusions.....	68

Appendices.....	69
Appendix A – Journal submitted paper	69
Appendix B – ECPA conference paper	85
Appendix C - Experiment map and schedule	97
Appendix D - Data sets.....	98
Appendix E - Daily data set.....	99
Appendix F – Normalization and Standardization	100
Appendix G – Active contour algorithm	101
Appendix H – Active contour iterations	102
Appendix I – Images after edge detection algorithm.....	103
Appendix J – Models.....	107
Appendix K – Statistics.....	108
Appendix L - Stepwise logistic regression.....	109
Appendix M – Confusion matrixes and ROC curves.....	111
Appendix N – F-test and T-test for classification	113
Appendix O – F-test and T-test for best time acquisition	114
References.....	115
Written References.....	115
Online References.....	120

List of Tables

Table 1. Applications of thermal imaging in the agriculture sector	29
Table 2. Hours of rounds and number of samples in each round	40
Table 3. Feature description.....	42
Table 4. Confusion matrix	45
Table 5. Point-biserial correlations.....	55
Table 6. All results from all models based on all data	56
Table 7. Results by day after infection.....	56
Table 8. Results after balancing healthy and infected samples (top), and with this balanced sample after selecting new features (bottom) (Exp1a).....	58
Table 9. Results after balancing between the infected days (top), and with this balanced data after selecting new features (bottom) (Exp1b)	59
Table 10. Results of the test set (Exp2a)	59
Table 11. Results of the test set (experiment No. 10446) (Exp2b)	60
Table 12. Results of a dataset with days 0,1,2 after infection (top), and the results of this same dataset with new features (bottom) (Exp3a)	61
Table 13. Results of a dataset with days 0,4,5,6,7 after infection (top), and the results of this same dataset with new features (bottom) (Exp3b)	61
Table 14. Results of ordinary regression (Exp4)	62
Table 15. A summary of all approaches and their results.....	65
Table 16. Performance for each round of the diurnal measurements	67
Table 17. Results of a dataset from 10:40 to 11:30 and after a new feature selection	67

List of Figures

Figure 1. Infrared on the electromagnetic spectrum (Ceramicx, 2011)	13
Figure 2. Lesions from downy mildew on the grapevine leaf surface (Gessler et al., 2011)	33
Figure 3. Flowchart of the research steps.....	37
Figure 4. Thermal image from first-round experiment.....	38
Figure 5. Examples of RGB images (top row) and thermal images (bottom row) of leaves from different infected days	40
Figure 6. Thermal images from infected and healthy leaves acquired at different times of the day ...	41
Figure 7. Prevalence of pixels of the entire image	47
Figure 8. The output of the active contour algorithm. Includes image, initial and final mask (A) and image with final mask (B)	48
Figure 9. Histograms (left) and boxplots (right) for each variable (a-m)	52
Figure 10. All features box-plots.....	52
Figure 11. All features box-plots without outliers.....	52
Figure 12. Heatmap of Pearson Correlation Coefficient Matrix	53
Figure 13. Pearson's partial correlations.....	54
Figure 14. Diagram representing the logical flow of the different analyses	57
Figure 15. Confusion matrix of ordinal regression	62
Figure 16. The first two rows of images have been correctly classified. The first row shows the masks of the Active Contour algorithm, resulting from the image in the second row. The same is true for rows three and four, but these are examples of incorrect classification.	63
Figure 17. Histograms of each selected variable:	67

1. Introduction

1.1. Problem description

The world population is expected to reach 9.8 billion people by 2050 (Ford et al., 2019). Due to the combined effects of a larger population and a richer diet, global food demand is projected to increase by 70% (Ford et al., 2019). Since arable land resources are limited, the pressure on productive land increases, and the per capita arable land is projected to decline from about 0.23 ha to about 0.15 ha (Seelan et al., 2003). To feed this amount of people, crop yield and production efficiency should be increased while minimizing inputs. Precision agriculture (PA) aims to use information and data processing for optimal management decisions for each field unit (Seelan et al., 2003).

Plant diseases are a major cause of production losses and economic losses in the agriculture industry (Savary et al., 2012). Pathogens are responsible for direct yield losses of 20-40% of global agricultural productivity (Savary et al., 2012). In order to ensure sustainable agriculture, it is essential to monitor plant health to prevent disease spread with as little damage to crop production as possible. However, the main challenge is the difficulty in determining the physical, chemical and biological changes in plants before symptoms of infection appear (Cui et al., 2018).

Disease detection techniques can be classified into invasive and non-invasive methods. Invasive techniques involve destructive leaf sampling followed by chemical treatments after direct identification of the pathogen (Sankaran et al., 2010). Non-invasive techniques identify plant diseases by detecting the impact of the pathogen on the physiological plant response. Currently, the most promising non-invasive disease detection methods are sensors that measure temperature, reflectance, or fluorescence (Mahlein, 2016).

Leaf temperature is a rapid response variable that can reveal crop stresses before visible symptoms appear (Khanal et al., 2017). Stressed plants respond with physiologic protection mechanisms that lead to changes in leaf surface temperature (Mahlein, 2016). Infrared thermography (IRT) enables the evaluation of the plant temperature related to changes in water status and transpiration due to infections by pathogens. IRT is a technique based on the relationship between leaf stomatal closure or aperture and its surface temperature (Mahlein, 2016). Temperature differences within individual leaves, plants, and crops indicate the presence of disease in plants (Mahlein, 2016). Compared to optical, multispectral, and

hyperspectral sensors, thermal sensors have shown to be more effective at detecting disease-induced early modifications (Khanal et al., 2017).

Plasmopara viticola, also known as oomycete grape downy mildew (DM), is a serious pathogen of grapevines that spreads through extremely efficient asexual reproduction cycles (Kiefer et al., 2002). Originally from North America, DM was accidentally introduced to Europe at the end of the 19th century, where it caused extensive damage to the grape industry (Gessler et al., 2011). The optimal environmental conditions for disease spread are high humidity and moderate temperatures. At early stages of infection, DM causes an increase in transpiration rate and a decrease in leaf temperature. Further DM development causes the appearance of chlorotic and necrotic tissue, increased water loss and the inability of plant tissue to regulate stomatal opening (Calderón et al., 2014). Thermal imaging has been successful in detecting DM (Caro, 2014; Wen et al., 2019). However, IRT is often subject to environmental factors such as ambient temperature, sunlight, rainfall, or wind speed (Mahlein, 2016). Changes in environmental conditions may affect leaf temperature, making it difficult to differentiate it from a change caused by infection or disease (Grant et al., 2006). Alchanatis et al. (2010) found that for estimating and mapping water status variability of cotton, best results from thermal images were achieved at midday (12:00-14:00).

1.2.Objectives

This study aimed to develop an algorithm for early detection of grapevine downy mildew using thermal imaging. The specific objectives were to:

1. Extract features for classification based on temperature and image processing algorithms.
2. Develop classification models to classify between infected and healthy grapevine leaves.
3. Determine the best time along the day to acquire thermal images for downy mildew detection.

1.3. Thesis structure

This thesis begins with a literature review presented in chapter 2. The methodology of the research is depicted in chapter 3. In chapter 4, algorithms are described, including the edge detection algorithm (4.1) and classification models (4.2). Results and discussion are presented in chapter 5 followed by conclusions in chapter 6.

2. Literature Review

The review begins by describing the concept of precision agriculture (2.1). This is followed by a review of thermal sensing (2.2) and recent applications of thermal sensing in agriculture (2.3). Next, a review of common methods of diseases detection (2.4), followed by vegetational thermal behavior (2.5). This is then followed by a review of downy mildew in general, as well as a review of grapevine downy mildew and IRTs of it (2.6). The next section focuses on improving image capture conditions (2.7).

2.1. Precision agriculture

The world demand for crop production is rising due to three main reasons: human population growth, increasing meat and dairy consumption and biofuel consumption (Ray et al., 2013). The estimate is that by 2050, global agricultural production demand may increase by 60%-110% (Ray et al., 2013). In parallel, 5-10 million hectares of farmland are lost annually due to the development of cities and their expansion, soil erosion, loss of efficiency as agricultural land, etc. (Farsund et al., 2015). Maintaining food security is one of the main global challenges in the 21st century (Farsund et al., 2015). Decades ago, farms were quite small and it was easy to detect all within-field variations and take suitable management decisions for each field unit. The developments of tractors and commercial fertilizers helped farmers to improve field management, increase production and expand their farms. These growths raised a problem with the difficulty and complexity of managing such large fields; the field holders had less connection and knowledge about specific characteristics of each field. Precision agriculture (PA) is designed to enable farmers to manage large fields by providing with detailed knowledge about each unit (Pedersen & Lind, 2017).

Introducing site-specific modern technologies can help to address forthcoming challenges and trends in agricultural production that require high precision, and improve crop yield significantly. It allows higher accuracy of farm records, and provide information that leads to better in-field management decisions. Although technology has a big potential to help to deal with the global challenges, an integrated approach is necessary to ensure its use among farmers (Seelan et al., 2003).

Precision agriculture is a cyclic process of collecting data from the field, analyzing it, and formatting evaluations that allow the farmer to make better management decisions (Pedersen & Lind, 2017). PA includes three major goals. First, increasing the profitability and

continuity of agricultural activities by optimizing the utilization of accessible resources. Second, reducing environmental footprints. Third, improving the quality of the agricultural work environment. Using information and communications technologies such as sensors, information systems, and enhanced machinery helps identify the variability of crop and soil conditions in the field (Gebbers & Adamchuk, 2010).

2.2. Thermal sensing

An object emits thermal radiation when its temperature is above absolute zero (-273.15°C). The higher the temperature, the larger the amount of thermal radiation emitted (Bulanon et al., 2008). The maximum level of radiant power an object emits depends only on its temperature (Vollmer & Möllmann, 2017). Infrared (IR) radiation, is a region of the electromagnetic (EM) radiation spectrum, which is divided into wavelength regions: visible (VIS; wavelength range, 400– 700 nm), near-infrared (NIR; 700– 1,100 nm), and shortwave infrared (SWIR; 1,100– 2,500 nm), and the energy emitted in the thermal infrared (TIR; 3 to 15 μm) (Martinelli et al., 2015). The electromagnetic spectrum provides a classification of the entire range of electromagnetic radiation according to their wavelengths, extending from the shortest cosmic rays to the longest broadcast band waves ([Figure 1](#)).

IR thermal imaging, also called thermography, is a measurement technique, which enables measurement of quantitatively object's surface temperatures. A way to detect emitted infrared radiation in the thermal infrared range from 8 to 12 μm is by thermographic and infrared cameras and is illustrated in false-color images, where each image pixel contains the temperature value of the measured object (Mahlein, Oerke, Steiner, & Dehne, 2012). The object emits radiation into the camera's direction, which focuses the IR waves to the detector and this is quantitatively measured. Still, this requires the IR radiation to pass through the space between the object surface and the camera detector.

The object's temperature, its material properties, and its surroundings influence the amount and distribution of emitted radiation as a function of wavelength (Vollmer & Möllmann, 2017). First, hot objects and the atmosphere can emit additional radiation to the camera lens which may impair the temperature reliability the detector has absorbed. Second, when passing through the atmosphere, IR windows or camera optics the radiation level might reduce (Vollmer & Möllmann, 2017). Thermal imaging technology enables us to measure temperature by non-contact, non-invasive, non-destructive, and high-resolution

measurements (Bulanon et al., 2008). Thermographic camera's performance is distinguished by three parameters: image resolution (image pixel size, number of image pixels, depth of focus), thermal sensitivity (thermal range, measurement precision), and scan speed (Mahlein, et al., 2012).

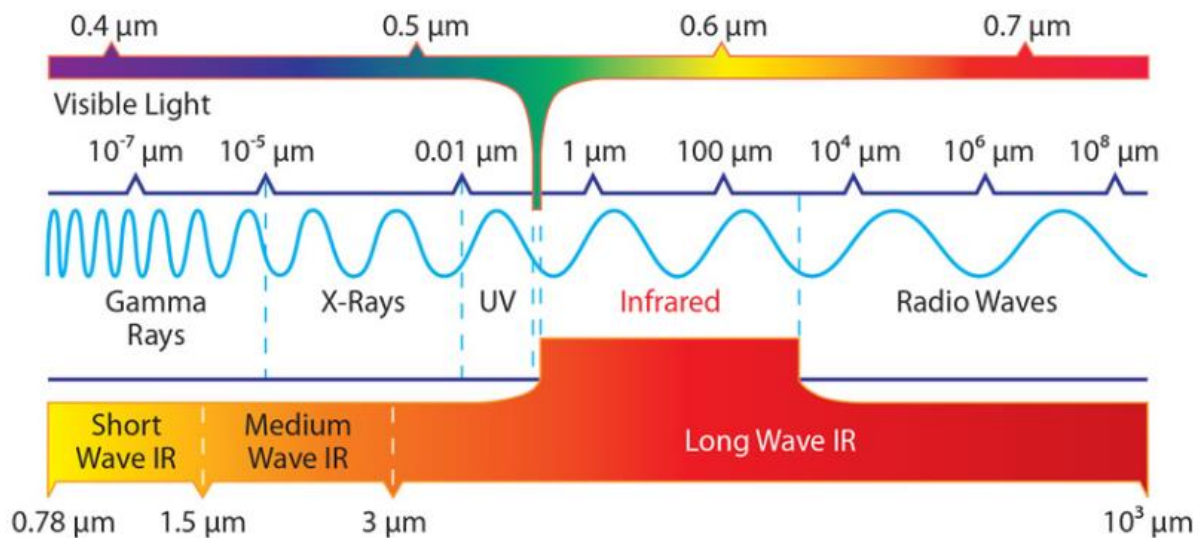


Figure 1. Infrared on the electromagnetic spectrum (Ceramicx, 2011)

2.3. Thermal sensing in agriculture

The plant's leaf temperature depends on internal and external physiological factors (Oerke et al., 2006). Water is the main source of infrared absorption in plant tissue and environmental factors such as solar radiation, air temperature, and relative humidity, influence the temperature of plants. Therefore, there is a correlation between water status and leaf temperature (Oerke et al., 2006). Thermal sensing enables the imaging of surface temperature differences of leaves, plants or crop canopies and can be applied in near or remote sensing. By thermographic cameras, it is possible to detect the emitted radiation in the thermal infrared (TIR) from 8– 12 μm and present it by image pixels (Mahlein et al., 2012).

Thermal sensing can be used in many aspects of crop and soil monitoring in the agricultural sector including estimating soil and crop water stress for irrigation scheduling, evaluating moisture stress, determining disease and pathogen-infected crops, estimating of residue cover, locating tiles in fields, monitoring crop maturity for harvesting, mapping and estimating crop yield. The following table ([Table 1](#)) surveys the recent applications of

thermal imaging and reveals that eight of them used thermal imaging, six used both thermal and multispectral imaging, six used both thermal and RGB imaging, three used both thermal imaging and sensors, and five used only thermal sensors.

Sector	Goal	Conclusions	Plants	Environment	Platform	Camera	Data	Features	Algorithm	Data Analysis	Reference
Soil moisture	Map the spatial distribution of soil moisture conditions and monitor its change over time.	Thermal images can be used for operational monitoring of soil moisture conditions at the field or regional scales. This means the PSMI is more straightforward and less time-consuming than those that use Ts to estimate soil moisture content.		field	satellite	Landsat medium-resolution satellite imagery, 60 m for Landsat-7 and 100 m for Landsat-8 and was re-sampled to 30 m and TIRS	soil moisture was measured in 18 agricultural fields	PSMI- Perpendicular Soil Moisture Index, Tmax, Tmin, Ts	A simple linear regression, ENVI image processing, MATLAB	R ² , RMSE	Shafian & Maas, 2015
Soil moisture	Estimates of surface soil moisture for a large field irrigated by a center pivot sprinkler system.	The significance of thermal imagery as the most relevant information in surface soil moisture estimations. A model with just thermal imagery can estimate the soil moisture values with RMSE of approximately 3%.	alfalfa and oats	field	UAV	Optical cameras at 0.15 m resolution and thermal camera. The wavelength range peaks around 420, 500, 600 and 800 nm. Multispectral sensor.	area of approximately 84 acres	red, blue, NIR, thermal, NDVI, VCI, EVI, and field capacity	Artificial Neural Network model, a K-fold cross validation when K=5, a wrapper selection method	R ² , RMSE, MAE, coefficient of correlation (r), coefficient of performance (e)	Hassan-Esfahani, Torres-Rua, Jensen, & McKee, 2015
Soil moisture	Evaluate the strength of relationships between soil moisture, mechanical resistance and thermal inertia calculated from the drop of surface temperature.	Coarse resolution airborne thermal images performed better than a handheld thermal gun There is significant relationship between thermal inertia and soil moisture and mechanical resistance.	grapes	grass covered vineyard	Airborne, handheld thermal gun	Thermal camera- FLIR model ThermoCAM SC2000. Handheld thermal gun with $\pm 1\%$ accuracy, TIR sensors - 8 to 14 μm	25 plots (4 m ² , each)- the model was based on 100 randomly sampled clusters with 5 replicates (total of 500)	Mechanical resistance and thermal inertia (Tlc)	Multiple linear regression, linear mixed-effects. FLIR analysis software, statistical software R	Pearson's correlation coefficients, AIC, RMSE, R ²	Soliman, Heck, Brenning, Brown, & Miller, 2013

Sector	Goal	Conclusions	Plants	Environment	Platform	Camera	Data	Features	Algorithm	Data Analysis	Reference
Soil surface wetness	An overview of the 'triangle' method for estimating soil surface wetness and evapotranspiration fraction.	The triangle method has the capability of producing non-linear solutions for surface moisture availability and surface evapotranspiration from large image data sets.			satellite			Tmin, Tmax, $\Delta T = T_{ir} - T_w$, NDVI	triangle method	RMSE, R^2	Carlson, 2007
Moisture stress	To estimate LAI by thermal imaging of the wheat crop canopy under different moisture stress conditions and to compare these estimates using thermal and digital imaging with those measured from a plant canopy analyzer.	Thermal image analysis can be applied as a non-destructive, rapid technique to characterize the temperature of crop canopy and then to estimate the LAI of wheat grown under moisture stress conditions. The support vector machine method gave the best result.	wheat	field	1 m above the canopy	Thermal and digital handheld camera- Testo 890-1 which supports a wide-angle lens of 42x32 and detector of 640x480 pixels. Can estimate temperatures from -20 to 100 C. FPA type of detector.	Field-77°12' E longitude, 28°35' N latitude, and an altitude of 228.16 m over the mean sea level	LAI	support vector machine (SVM), the maximum likelihood method, Mahalanobis algorithm, Minimum distance to mean, Parallelepiped	R^2 , RMSE, MBE, ANOVA, confusion or error matrix	Banerjee, Krishnan, & Mridha, 2018

Sector	Goal	Conclusions	Plants	Environment	Platform	Camera	Data	Features	Algorithm	Data Analysis	Reference
Crop water stress	Developing an adaptive algorithm capable of detecting erroneous irrigation signals or limiting water delivery under low atmospheric demands.	The new algorithm CWSI-DT better than NP. It is concluded that the crop water stress index can become more efficient in conjunction with a well-developed control algorithm.	apple	orchard	ground	Shielded air temperature sensors- Model 109, Campbell Scientific. Infrared thermometers- IRT/c.2: Type J, Exergen.	orchard - coordinates of latitude 46.26°N, longitude 119.74°W, and 360 m above sea level	CWSI, Tc, air temperature Ta	linear regression, CWSI-DT algorithm	MAE, RMSE, standard deviation and standard error of mean, R ²	Osroosh, Troy Peters, Campbell, & Zhang, 2015
Crop water stress	The assessment of heterogeneity in water status in a commercial orchard, as a prerequisite for precision irrigation management.	Indicators derived from thermal imagery acquired from an UAV over a commercial orchard have been shown to accurately describe the spatial variability in crop water status.	almond, apricot, peach, lemon and orange	orchard	UAV	MIRACLE 307, Thermoteknix Systems- spectral region 8–12 µm, resolution of 640 x 480 pixels with a field of view of 458	42 ha orchard	SWP, CWSI, Tc-Ta	regression	Kolmogorov–Smirnov, Skewness and Kurtosis indices with the SPSS	Gonzalez-Dugo et al., 2013
Crop water stress	Assessment if the CWSI-TT method can be an effective trigger for automatically scheduling either full or deficit in a semi-arid region.	CWSI-TT is an effective method for irrigation scheduling of grain sorghum. A shortcoming of using this method is that false-positive irrigation triggers may be generated early in the season, which could lead to over-irrigation.	Sorghum	field	1.5 m above the crop canopy	infrared thermometers (IRTs)- Exergen model IRT/c.5:1- Type T	10 m ² area in each of the 48 treatment plots	NWSB, CWSI, CWSI-TT, RDI, LWP, Ta	General Linear Models	Bonferroni t-test Statistical software was SAS	O'Shaughnessy , Evett, Colaizzi, & Howell, 2012

Sector	Goal	Conclusions	Plants	Environment	Platform	Camera	Data	Features	Algorithm	Data Analysis	Reference
Crop water stress	Estimates of the effectiveness of using a TTT method for automatic irrigation scheduling of cotton irrigated with LEPA drag socks under a center pivot system.	The TTT method, when applied automatically, was successful in producing yields and water use efficiencies as good as those resulting from scientific irrigation scheduling using the neutron probe, and better irrigation water use efficiencies than achieved using the neutron probe.	cotton	field		infrared thermometer sensors (IRTs)- Exergen model IRT/c.5:1-Type T	18- row plots, 10 m ² area in each plot	SWD, ET, WUE, IWUE	TTT algorithm	SAS Proc Mixed models	S. A. O'Shaughnessy & Evett, 2010
Crop water stress	To estimate evapotranspiration fluxes over a semi-arid agricultural area.	Estimation of ET within wide spatial scales by one-layer models and integration of ground-based meteorological data with satellite observations is a useful tool for quantifying and controlling water consumption.	orange	orchard	satellite	Landsat Thematic Mapper TM5	area about 18,000 ha	Ts-Ta, EF, CWSI, NDVI, Kc	linear transformation	R ²	Barbagallo, Consoli, & Russo, 2009
Crop water stress	To assess the reliability of CWSI for detecting water stress.	The CWSI, obtained from continuous nadir-view measurements with IRTs, is a good and very sensitive indicator of water stress in pistachio.	pistachio	orchard	canopy top	4 IRTs- model IRR-P, Precision Infra-Red Radiometer with angular field of view is 44°.	Two large plots of 32 ha each	CWSI, NWSB, Tc-Ta, VPD	linear regression	R ² , coefficient of variation Cv (the ratio of the standard deviation to the mean)	Testi, Goldhamer, Iniesta, & Salinas, 2008

Sector	Goal	Conclusions	Plants	Environment	Platform	Camera	Data	Features	Algorithm	Data Analysis	Reference
Crop water stress	Simplify the calculation of the crop water stress index (CWSI) and improve its diagnostic accuracy.	1. Canny algorithm can well eliminate soil pixels from UAV high resolution thermal infrared images; 2. CWSIs calculated using Twet, Tdry, and TI is simpler and more robust than CWSIe, CWSIs, NDVI, TCARI, OSAVI; 3. soil volumetric water content at different depths and stomatal conductance are highly correlated with CWSIsi.	cotton	field	UAV	Thermal infrared camera- Zenmuse XT, FLIR System with spatial resolution of 640×512 pixels, a spectral range of 7.5–13.5 μm and multispectral sensor u-MCA, Tetracam	four irrigation treatment plots	TI, Twet, Tdry	Canny edge detection algorithms, Expectation–Maximization algorithm	R ² , Pearson's correlation coefficient- R	Bian et al., 2019
Crop water stress	Define the methodology to assess the almond water status by means of thermal information.	1. CWSI is the most appropriate thermal index to monitor the water status. 2. The row-level images were enough robust to estimate the water status.	almond	orchard	Row level - 1.5m of height. tree level- Ground	ThermaCam- Flir SC660, Flir Systems, 7–13 μm, 640×480 pixels	four replications- each had 15 trees (3 rows and 5 trees per row).	TC, ΔT canopy-air, CWSI, IG, Ψleaf, gs		Levene's test to check the variance homogeneity of the studied variables	García-Tejero et al., 2018

Sector	Goal	Conclusions	Plants	Environment	Platform	Camera	Data	Features	Algorithm	Data Analysis	Reference
Crop water stress	Develop a new system involving thermal imaging and machine learning to assess the vineyard water status; and to provide full trained, robust prediction models.	Using the models trained with reference temperatures demonstrated that canopy temperatures, along with Tdry and Twet and weather parameters, would be very suitable for the useful prediction of the vineyard water status.	grapevine	vineyard	1.2 m of distance and 1.0 m from the ground	FLIR A35, FLIR1 Systems	four replications, for each 25 plants - 15 middle ones were the ones in which the measurements were taken.	VPD, Ψ_{stem} , Twet, Tdry, CWSI, Ig	regression models	RMSE, R ²	Gutiérrez et al., 2018
Crop water stress	evaluate the possibility of using small unmanned aerial system (UAS)-based remote sensing technologies to monitor the crop stress	1. GNDVI, CC and CT were able to differentiate crops with full and deficit irrigation treatments at each of the three growth stages. 2. Performance of small UAS- based imagery-based indicators in crop stress monitoring and crop yield estimation was better than or comparable to that of the ground-based LAI estimates.	pinto beans	research farm	UAS, ground	Multispectral digital camera with single optical lens- NiteCanon ELPH110 LDP LLC and infrared thermal camera- Tau 640 uncooled cores, Flir Systems.	Four irrigations plots of 80.1 m by 12.2 m in dimension. Each irrigation plot was split into four tillage plots of 18.3 m by 12.2 m.	GNDVI, DNI, CC, CT, LAI	linear regression	ANOVA ('PROC GLM'), 'PROC CORR' procedure with 'Pearson'- r	Zhou et al., 2018

Sector	Goal	Conclusions	Plants	Environment	Platform	Camera	Data	Features	Algorithm	Data Analysis	Reference
Evapotranspiration (ET) and drought stress	To compare the evaporative stress index (ESI) derived within a thermal remote sensing energy balance framework, with prognostic estimates of soil moisture (SM), ET, and runoff anomalies generated with the NLDAS.	Of the remote sensing indices evaluated, ESI was best correlated with NLDAS indices and with USDM classes. Both ESI and related remote sensing ET index ETI outperformed anomalies in NLDAS daily ET. Both ESI and NLDAS SM change products indicated value in providing early warning of changing drought conditions recorded in the USDM.	corn and soybean		satellite	TIR-based remote sensing Atmosphere–Land Exchange Inverse (ALEXI) model and NOAA-AVHRR sensor series		SM, ET, VHI, SPI	smoothing algorithm, Atmosphere–Land Exchange Inverse (ALEXI) model	Pearson correlation coefficient, Spearman nonparametric coefficient	Anderson et al., 2013

Sector	Goal	Conclusions	Plants	Environment	Platform	Camera	Data	Features	Algorithm	Data Analysis	Reference
Monitoring crops	Develop fully automat crop monitoring for moving irrigation systems based on thermal-RGB imagery, and develop an image-processing algorithm to extract canopy surface temperature and percent canopy coverage from images.	The amended image-processing algorithm was able to successfully extract canopy surface temperature and percent canopy cover. The current design of imager allows for creating a network of imaging units in the field to obtain real-time surface temperature data from plant canopies. The first peak of thermal image relative frequency histogram could be a more accurate representative of sunlit leaf surface temperature.	mint	field	ground	thermal module with shutter (FLIR Lepton® 2.5) with 80X60 pixels, 8–14 µm, accuracy of ±5°C. RGB Raspberry Pi camera module (V2, Raspberry PiFoundation), with 3280×2464 pixels	The area was approximately 18 m ²	surface temperature	a linear regression	RMSE, MAE, SE of mean, R ²	Osroosh, Khot, & Peters, 2018
Plant disease detection	Detection, identification and quantification of plant diseases on different scales.	IR Airborne thermography is highly suitable for the detection of disease-induced modifications in plant transpiration and water status. Common instruments are active sensor systems; however, they are limited by the response time.						temperature differences between leaf tissue and air temperature, MTD, SVI	ANN, SVM, PCA, SAM		Mahlein et al., 2012

Sector	Goal	Conclusions	Plants	Environment	Platform	Camera	Data	Features	Algorithm	Data Analysis	Reference
Plant disease detection	Innovative and reliable alternatives for objective and reproducible detection and monitoring of plant diseases are demanded.	Successfully combine the information from thermography, chlorophyll fluorescence and spectral reflectance for an early detection of plant diseases. An approach with different non-invasive sensors provides a more comprehensive view on plant physiology during pathogenesis than each sensor alone.	cucumber	field or greenhouse	non-invasive sensors	IR scanning camera- VARIOSCAN 3201. Spectral sensitivity 8-12 μm , a geometric resolution of 1x5 m radians. Hyperspectral camera- ImSpector V10E.	Four treatments- 10 plants each, five plants for optical sensor measurement and five plants for destructive analysis.	MTD, NDVI, ARI	general linear model, Tukey's test, Wilk's lambda test. performed using SPSS v. 17.0.		Berdugo, Zito, Paulus, & Mahlein, 2014
Plant disease detection	Explored the use of high-resolution thermal and multi-spectral imagery as an indicator of DM infection.	Results demonstrated that canopy temperature and the green/red ratio were relate to physiological stress caused by DM infection at high NDVI steps. In other words, enabled the identification of infected plants by using these indices.	poppy	field	UAV	3-band multi-spectral camera- ADC Lite, Tetracam. 640x480 pixel thermal camera - MIRICLE 307, Thermoteknix with 14.25 mm f1.3 lens.	Two experiments were conducted in two fields (45 x 20 m), approximately 25 plants per m^2	average, minimum and maximum temperatures, Tc-Ta, R550/R670, NDVI	general linear model, grid	ANOVA test	Calderón, Montes-Borrego, Landa, Navas-Cortés, & Zarco-Tejada, 2014

Sector	Goal	Conclusions	Plants	Environment	Platform	Camera	Data	Features	Algorithm	Data Analysis	Reference
Plant disease detection	sensing and quantifying apple scab	Thermographic measurements revealed differences in disease severity. MTD is robust to side-effects from abiotic factors. Thermal infrared imaging has great potential as a sensor technique for precision agriculture but the technique has to be combined with other sensors.	apple	glasshouse	ground	Varioscan 3201 ST. spectral sensitivity 8-12 μm , a geometric resolution of 1.5 mrad and a 30°x20° field of view lens. Digital camera- Nikon D50		MTD, minimum, maximum and mean leaf temperature, standard deviation	distinct leaf areas of lower temperature precede the appearance of scab	ANOVA, HSD test, Fisher's LSD test, t test statistics, Pearson's coefficient of determination (R^2)	Oerke, Fröhling, & Steiner, 2011

Sector	Goal	Conclusions	Plants	Environment	Platform	Camera	Data	Features	Algorithm	Data Analysis	Reference
Plant disease detection	1. To evaluate the relationship between disease severity and MTD of cucumber leaves infected. 2. To assess the impact of environmental conditions to describe the potential for the assessment and quantification of downy mildew in the field.	Changes in temperature of infected leaves allowed the discrimination between healthy and infected areas in thermograms, even before visible symptoms of downy mildew appeared. This remote sensing technique interferes as little as possible with the plant. There was a strong linear correlation between the percentage infected leaf area and MTD but MTD alone is not suitable for the quantification of downy mildew severity in the field.	cucumber	greenhouse	ground	Thermal camera - VARIOSCAN 3201 ST- Spectral sensitivity 8- 12 μm and a geometric resolution of 1.5 mrad. Thermal resolution 0.03 K, accuracy less than ± 2 K. Digital camera- JD 4100 Z3, Jenoptic.		MTD, air temperature, humidity, water content, age of the leaf, mean leaf temperature, variance.	linear and quadratic regression	ANOVA, Pearson coefficient (r), F-values, mean comparisons	Oerke et al., 2006

Sector	Goal	Conclusions	Plants	Environment	Platform	Camera	Data	Features	Algorithm	Data Analysis	Reference
Plant disease detection	Explore the use of thermal imagery, chlorophyll fluorescence, structural and physiological indices calculated from multispectral and hyperspectral imagery as early indicators of water stress caused by VW infection and severity.	The results indicate that the CWSI obtained from high spatial resolution thermal imagery can be used to detect the lower transpiration rates induced by V. dahliae infection. CWSI values estimated were significantly lower for healthy trees than for those affected by the disease. Also, Tc – Ta and G allowed identifying trees at the early stages of disease development.	olive	orchards naturally	UAV	Thermal camera- MIRICLE 307, Thermoteknix. 14.25 mm f1.3 lens, spectral range 8-12 µm. 640 ×480-pixel resolution. 6-band multispectral camera- MCA-6, Tetracam. Hyperspectral imager Micro-Hyperspect VNIR model.	Two study sites. The first in a 7-ha orchard; trees were planted at a spacing of 6×4 m. The second in a 10-ha orchard; trees were planted at a spacing of 6×3 m.	Tc, Ta, G, PRI570, PRI, crown temperature, CWSI, B, BG1, BR1, FLD3, NDVI, PRI515, HI, chlorophyll and carotenoid	linear model (GLM) procedure of SAS 9.2	ANOVA, Dunnett's two-tailed test.	Calderón, Navas-Cortés, Lucena, & Zarco-Tejada, 2013
Plant disease detection	To test whether infrared thermography can be used to distinguish between infected versus non-infected areas upon attack by P. viticola within single grapevine leaves under varying water status.	Thermal imagery clearly perceived the beginning of an infection. Using the high sensitivity to temporal and spatial variation in leaf temperature, ground based remote-sensing may have important applications in disease forecast models or other screening activities and technical applications.	vineyards	greenhouse and field	ground	TH7102 MX-waveband of 8–14 µm. Geometric resolution of 1.58 m radians. The thermal resolution is 0.06°C and the accuracy of absolute temperature measurement less than ±2°C.	The vineyard was planted with a vine density of 2.0 x 1.0 m	MTD		Two Way ANOVA, Fisher's test	Stoll, Schultz, Baecker, & Berkelmann-Loehnertz, 2008

Sector	Goal	Conclusions	Plants	Environment	Platform	Camera	Data	Features	Algorithm	Data Analysis	Reference
Plant disease detection	Early identification of emerging biotic stresses in crops, facilitating the containment of disease outbreaks.	Non-destructive imaging techniques, chlorophyll fluorescence imaging and thermography were able to visualize infections at an early stage before damage appeared. However, the early phase of the attack is not characterized by a rise in temperature, but rather by immediate cooling.	sugar beet	growth-room	above the same plant leaf region	thermal camera FLIR\Agema THV900LW - pixel resolution 272×136. Video camera- Watec colour CCD, LCL-217HS. Fluorescence imaging system		Chlorophyll fluorescence, PEA			Chaerle, Hagenbeek, De Bruyne, Valcke, & Van Der Straeten, 2004
Residue cover and tillage mapping	The objective of this study was to evaluate high spectral resolution remote sensing (RS) data for rapid quantification of residue cover.	Airborne thermal IR imagery can be used for crop residue variability assessment. VIS and NIR ATLAS datasets more clearly differentiated among treatments compared with handheld spectroradiometer data. High spectral resolution TIR imagery more accurately delineated treatment differences compared with VIS and NIR spectra.	Wheat	Laboratory	handheld or airborne	Sensors- GER 1500 Spectroradiometer, Airborne Terrestrial Applications, and digital camera	Two experiments in two fields (45 x 20 m). The seeds were sown at a spacing of 80 mm x 0.50 m, approximately 25 plants per m ² .	Emitted energy, reflected energy (indirect)	stepwise linear regression, general linear model (GLM)	ANOVA, PCA	Sullivan et al., 2004

Sector	Goal	Conclusions	Plants	Environment	Platform	Camera	Data	Features	Algorithm	Data Analysis	Reference
Crop maturity mapping	Examine of employing thermography for the non-destructive determination of produce characteristics and the thereby resulting measures of quality maintenance and to fix the limits of its use.	Thermal imaging could be used for-analysis of pre-harvest conditions, evaluation of post-harvest climate, microbial infestation, freshness status of produce.	different kinds	laboratory	Ground	IR detector, type MCT- spectral range 8-12 μ m Temperature resolution <0.1 K. Measurement ranges between - 40°C and +1,200°C and for object distances as of 25 cm.		temperature difference, transpiration resistance values, surface temperatures			Linke, Beuche, Geyer, & Hellebrand, 2000
Crop yield mapping	Demonstrate the potential of using thermal imaging information to segment orange fruits from the canopy as aid to robotic fruit harvesting.	The fruits were successfully segmented in the thermal images using image processing techniques during the time range of the largest temperature difference, which suggest potential application of thermal imaging for improved detection for harvesting.	orange	orchard	Ground	Thermal camera- ThermaCAM P65HS with 320x240 pixels, spectral range 7.5-13 μ m, resolution of 0.05C and accuracy of +- 2%. Temperature ranges -40 to 120 C. Digital color camera D-490ZOOM, Olympus.		Surface temperature of the fruit, ambient temperature and relative humidity (RH)	segmentation approach using the histogram tail method, cluster separation algorithm-based thresholding		Bulanon et al., 2008

Sector	Goal	Conclusions	Plants	Environment	Platform	Camera	Data	Features	Algorithm	Data Analysis	Reference
Crop yield mapping	Estimating the number of apple fruits and measuring their diameter for calculating the current yield	The measuring technique based on thermal imaging and analysis can be employed to provide an objective and easy counting of apples and measurement of their diameters. The system was used successfully all cases when only a small part of the apple was detected. However, it was not always able to distinguish between fruits and leaves growing deep in the tree-crown.	apple	orchard	Ground	Thermal camera AGEMA 570, Flir SystemsTM	A thermal camera captured images of apple trees five times during the vegetation period. Each time 120 images of twenty apple trees.	Temperature gradient between fruits and their background, average temperature, NDI	a fruit detection algorithm	Correlation coefficients R ²	Stajanko, Lakota, & Hočevár, 2004
Crop yield counting	Propose an automatic wheat ear counting system using thermal images.	The algorithm did demonstrate high correlations with various manual image-based ear counts. There should be a difference of at least 2 °C between the ear and leaf temperatures for this thermal ear counting algorithm to work.	wheat	field	Handheld, above the canopy	MIDAS 320L IR camera, DIAS IR GmbH with a -20C to 120°C temperature range, 8–14 µm spectral range in one channel, 320×240 radiometric detector. RGB images- Sony QX1-ILCE camera.	72 plots per trial (3 replicates × 24 genotypes) The genotypes were evaluated in 9 m ² , 6 rows, 0.25 m apart and a planting density of 250 seeds per m ² .	Temperature differential between the ears and the rest of the canopy.	Linear regression	R ²	Fernandez-Gallego et al., 2019

Table 1. Applications of thermal imaging in the agriculture sector

2.4. Disease detection

Diseases in plants are caused by is biophysical and biochemical changes in host plants due to fungal or fungal-like organisms, bacteria, viruses, and viroid. They affect the quality of fruits, vegetables, grains, legumes, and can lead to mortality in plants (Berdugo et al., 2014; Golhani, Balasundram, Vadamalai, & Pradhan, 2018). These diseases become a major threat to global food security and lead to significant production and economic losses in the agricultural industry (Golhani et al., 2018). Each year 10%–16% of losses in the global harvest of crops causes by plant diseases and costing an estimated by US\$220 billion. Therefore, it is critical for sustainable agriculture to monitor plant health to prevent disease spread with the least damage to crop production and to develop effective early disease detection techniques. Early information on disease development and the quantity of the disease can enable disease control using suitable management strategies such as pesticides according to crop protection applications, fungicide applications, and disease-specific chemical applications (Sankaran et al., 2010). However, the main challenge is the difficulty in determining the physical, chemical, and biological changes in plants early, before there are symptoms of an infection. The extent of pathogen development in a plant is not always correlated positively with disease intensity, since the amount of pathogen biomass is not necessarily proportional to the extent of visible disease symptoms. Another challenge is the costs and time associated with these monitoring techniques that must be continuously executed (Cui et al., 2018).

Prevalent methods for the diagnosis and detection of plant diseases include visual plant disease estimation by human raters, microscopic evaluation of morphology features to identify pathogens, as well as molecular, serological, and microbiological diagnostic techniques (Mahlein, 2016). Disease detection techniques can be classified into invasive and non-invasive methods (Sankaran et al., 2010). Invasive techniques are used after the onset of plant disease symptoms and involve destructive leaf sampling followed by chemical treatments (Sankaran et al., 2010). Techniques such as Western blotting, Enzyme-Linked Immuno-Sorbent Assay (ELISA), Reverse Transcription Polymerase Chain Reaction (RT-PCR), and microarrays. These techniques are speedy and accurate in disease detection. Still, they pose inconsistency and insensitivity due to different reasons, including host-pathogen interaction and quantify disease concentration (Golhani et al., 2018).

Non-invasive techniques identify plant diseases by detecting the impact of the pathogen on the physiological plant response and not through direct identification of the pathogen. Techniques such as fluorescence spectroscopy, Visible/Near-Infrared (VNIR) spectroscopy, fluorescence imaging, and hyperspectral imaging, detect morphological changes, transpiration rate changes, and volatile organic compounds (VOCs) profiles (Cui et al., 2018). They are reliable at early (asymptomatic) stages and rapid. However, some modifications and improvements are still needed (Cui et al., 2018).

2.5. Vegetational thermal behavior

Plant temperature is determined by loss of water through specialized pores in the leaves, called stomatal transpiration (Oerke et al., 2006). When leaf transpires, water is lost through stomata and leaf temperature decreases. However, if stomatal transpiration stops, leaf temperature increases as no heat dissipation occurs. Also, significant modifications in leaf temperature are caused by a pathogenic organism's influence on the conductance of plant tissue. Visible symptoms of diseases often appear only after latent settlement in the plant tissue; therefore, those organisms may be detected only by their effect on plant tissue. As leaf temperature is directly related to the rate of water evapotranspiration, infrared sensing has the potential for early disease detection, since leaf temperature enables the quantitative analysis of spatial and dynamic physiological information on the plant status without interfering with plants. A useful parameter for the differentiation of infected and non-infected tissue is the maximum temperature difference (MTD) within a leaf (Oerke et al., 2006).

Vegetation spectral signature is based on non-visible wavelengths that are reflected or absorbed (Vollmer & Möllmann, 2017). The spectral signature is influenced by the leaf pigments, cell structure, and water content. Healthy plants absorb properly the radiation in the visible region and reflect properly the radiation in the NIR region. Plants that are stressed or infected can be identified by their spectral signatures that are different from the signature of healthy, unstressed vegetation (Ashourloo et al., 2014). A plant that is in stressed conditions reacts with protection mechanisms that cause changes in variables such as surface temperature, or chlorophyll content that allows the plant to produce food for itself from a combination of water, carbon dioxide, and sunlight. At the early stages, when visual symptomless to the existence of a pathogen, plants react with physiological protection

mechanisms such as the reduction of the photosynthesis rate, which induces an increase of heat emission and fluorescence (the emission of light). At a later stage, pathogens cause a reduction of chlorophyll content that increases reflectance in the VIS and causes a shift of the red- edge position (the region of rapid change in reflectance of vegetation in the NIR range of the electromagnetic spectrum). Finally, these changes lead to suboptimal growth, change the thermal properties which in turn influence the radiation emitted in the TIR domain of the spectrum and producing a different spectral signature (Martinelli et al., 2015).

2.6. Downy mildew

Downy mildew (DM) is a devastating disease caused by the biotrophic obligate Oomycete *Peronospora* (Calderón et al., 2014). Oomycetes are a diverse group of organisms widely distributed in very diverse environments. Most oomycete species are parasites of plants. Evolution processes that occurred in different lineages of the oomycote have resulted in obligate biotrophic pathogens such as downy mildew. All species of downy mildew are obligate parasites with limited host range and low tissue damage of the infected plants (Caro, 2014). The severity of *Peronospora* attacks strongly depends on the duration of optimal environmental conditions for disease development. High relative humidity and moderate temperature conditions are favored disease spread (Calderón et al., 2014). DM pathogens can cause local or systemic infections in plants, especially in the colonized plant tissue. Such infections result in changes in the metabolic processes of plant tissues including shifts in respiration, photosynthesis, and transpiration (Oerke et al., 2006). *Peronospora* can generate germ tubes and invade the leaves by a direct mode making use of appressoria and by an indirect mode through the stomata on the abaxial surface (Caro, 2014). Destruction of leaf cuticle, increased permeability of leaf cell membranes, or inhibition of stomatal closure can increase water loss from infected leaf areas. Stomatal closure, obstruction of xylem elements and stomata, and defoliation may cause a reduction of transpiration (Oerke et al., 2006). The first symptoms are small, chlorotic to light-yellow leaf lesions on the upper surface without loss of vitality in plant cells (Calderón et al., 2014). These symptoms can evolve to a second stage in which affected tissues are curled, thickened, and become deformed and necrotic as the disease develops (Oerke et al., 2006). Lesions expand in size with time and often coalesce, eventually giving rise to large necrotic and brown areas in leaves. Further development of lesions may result in the death of the entire leaf (Calderón et

al., 2014). Downy mildew caused by *Peronospora* is one of the most destructive and economically restrictive diseases worldwide, observed to produce asymptomatic infections and therefore difficult to control (Caro, 2014).

2.6.1. Grapevine downy mildew

Grapevine is susceptible to many pathogens (Aziz et al., 2003). Oomycete *Peronospora viticola* is the causal agent of grapevine downy mildew (Aziz et al., 2003). *Peronospora viticola* was first observed in the southeastern United States. Shortly after this first observation, the pathogen was introduced to European countries and since then has posed a major problem in vineyards all over the world (Kieffer et al., 2002). For example, in 1915, 70% of the French grape crop was lost because of *P. viticola*. In 1930, 20 million hl of wine were lost in France. From 1907 to 1916, downy mildew was responsible for a 33% reduction in the total vine-growing area in the Baden province of Germany. Significant damage was also reported in Italy (Gessler et al., 2011). Because of numbers and results like these, downy mildew has been considered the most devastating disease to affect European vineyards. At optimal environmental conditions, it can infect large areas within a short period of time and causes huge losses in yield by affecting leaves and young grapes (Kieffer et al., 2002). Using a calendar-based schedule of uniform fungicide spraying of vineyards in grape-growing areas is an efficient way to control this disease. Considering that primary infections occur in discrete foci rather than uniformly disseminating, it would be advantageous to detect those initial infection foci and operate targeted treatments instead of the current homogenous and unselective spraying (Oberti et al., 2014). The disease will first appear on the upper leaf surface as yellow spots referred to as oil spots ([Figure 2](#)). Its germination is most abundant in the early spring (Gessler et al., 2011).



Figure 2. Lesions from downy mildew on the grapevine leaf surface (Gessler et al., 2011)

2.6.2. Leaf temperature and downy mildew

The temperature of healthy leaves varies with leaf transpiration and, therefore, is a function of stomatal conductance (Lindenthal et al., 2005). One of the effects on stomatal conductance is by foliar pathogens which infect plants (Lindenthal et al., 2005). DM development causes the appearance of chlorotic and necrotic tissue, increased water loss, and the inability of plant tissue to regulate stomatal opening (Calderón et al., 2014). IR thermography was evaluated as a non-invasive method for detecting the infection (Caro, 2014). Studies using digital infrared thermography reported an increase in the highest temperature difference within a leaf with the formation of necrotic tissue. They found that this was related to disease severity and that it could be used to discriminate between healthy and infected areas in thermograms, even before the appearance of visible DM symptoms (Calderón et al., 2014). IRT was used to detect spread of rose downy mildew infections one or two days before the appearance of visible symptoms (Caro, 2014) and cucumber downy mildew before visual symptoms as well (Wen et al., 2019). The initial signs of infection in the thermal images were observed as early as 4 days after infection.

2.7. Image capture conditions

Several factors affect the characteristics of the images, such as the time of day, the position of the sun in relation to the leaf, and the amount of overcast (Barbedo, 2016). In addition, when monitoring plants over time, if environmental conditions change, then leaf temperature may also change. When leaf temperature changes due to increased stress (or the onset of stress), it can be difficult to differentiate it from a change resulting from increasing air temperature or other environmental changes. Moreover, the differences between the different infrared variables (e.g., T_{leaf} , CWSI) vary with date, making it difficult to determine which variable is most sensitive to plant stress (Grant et al., 2006). According to Jones (1999), the sensitivity of different variables using leaf temperature can be very different depending on a leaf's stomatal conductance. Ideally, all images should be captured under the same conditions. In practice, this can only be achieved in controlled conditions, such as laboratories. However, the use of algorithms for automatic disease identification is more efficient if it can be applied in the field, where the condition can be controlled partially, if at all. In this regard, a more realistic approach would be to study the impact of the main factors affecting segmentation and then design experiments accordingly (Barbedo,

2016). According to the study of Fuchs (1990), near solar noon, the highest temperatures and the greatest canopy temperature differences were recorded. Alchanatis et al. (2010) found that for estimating and mapping water status variability of cotton, best results from thermal images were achieved at midday (12:00-14:00). This might be also the best time to detect temperature differences between infected and healthy leaves. Under field measurements, however, even healthy plants may close their stomata before solar noon, particularly in hot, sunny conditions (Correia et al., 1990).

3. Methods

3.1. Overview

Two experiments were conducted on controlled greenhouse plants with leaves that were healthy or at various stages of the disease. The leaves were infected in a controlled manner and images were acquired using a thermal camera. The research involved several steps as described in [Figure 3](#).

Experiment A ([section 3.2.1](#)) involved six campaigns acquiring images at day 0 (healthy leaves) and one, two, four, five, six, and seven days after infection (infected leaves).

Experiment B ([section 3.2.2](#)) involved acquiring images at seven different times along the day of leaves at six days after infection to determine the best time for image acquisition

The best model to classify between healthy and infected leaves was determined through several steps. The thermal images were segmented using the 'Chan Vese' active contour algorithm. Features were extracted from the segmented leaves based on temperature measurements in the resulting mask. Additional features were extracted from meteorological measurements from data obtained from a meteorological station. Feature selection for classification was conducted using Pearson's correlation coefficient and stepwise logistic regression analysis. Five classification models were developed and compared. The best model and time for image acquisition were selected based on accuracy, F1 score, and AUC.

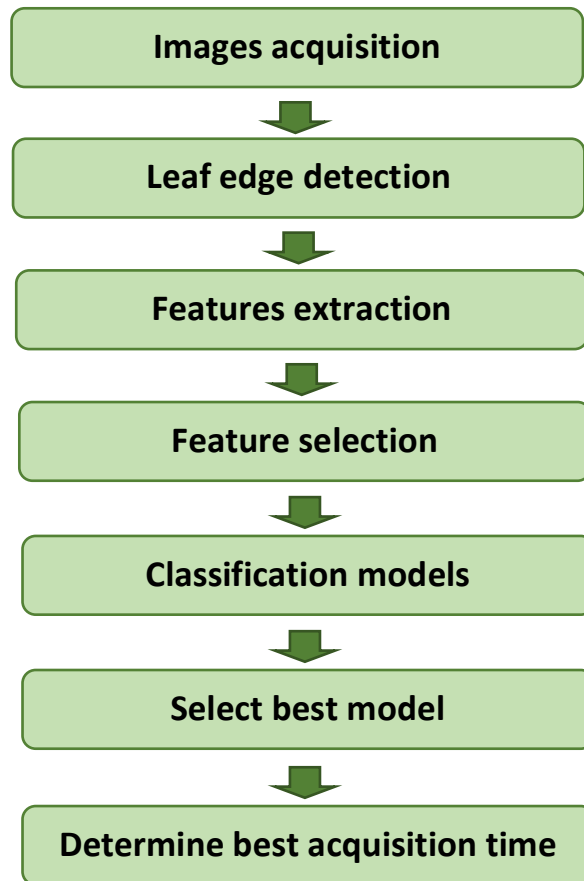


Figure 3. Flowchart of the research steps

3.2. Experimental design

3.2.1. Experiment A

Controlled experiments were conducted in six campaigns between the end of December 2019 and the end of October 2020 on 169 grapevine plants grown in experimental greenhouses in Evogene, Israel (31°47' 20.472" N, 35°12' 3.888" E). Each campaign lasted 12 days and included different imaging days (healthy leaves and 1, 2, 4, 5, 6, and 7 days after infection) in which around 30 plants were tested as detailed in [Appendix C](#). Plants received two daily doses of water. Each of the six campaigns included the following stages:

1. The leaves that were selected to be infected were marked by color clips or aluminum foil.
2. On the first day of the campaign, images of the healthy leaves were acquired.
3. 2-6 leaves in each plant were infected with downy mildew and were inserted in a high humidity chamber which allowed optimal conditions for the disease to develop.

4. Images of each infected leaf were acquired on 1, 2, 4, 5, 6 and 7 days after infection.
5. After the last imaging day (day number 7), the leaves were placed in Petri dishes to evaluate the disease severity.

To allow optimal photosynthesis of the plants, imaging was conducted outside of the greenhouse. The plants were taken out of the controlled greenhouse and placed outdoors for at least one hour before imaging to allow them to adjust to the environmental conditions that were different from greenhouse conditions. The images were acquired between 10 a.m. and 3 p.m. to ensure high solar radiation that allowed plants to conduct photosynthesis. Images were only acquired on sunny days without clouds.

Meteorological conditions were continuously monitored and included measurements of air temperature (°C), relative humidity, solar radiation (W/m²), wind speed (m/sec), and wind direction.

An image of each leaf was acquired by two cameras - a thermal camera and an RGB camera. For each leaf, the thermal camera acquired a half-minute video and the RGB camera acquired two images. One image from each video was selected for classification. The image was manually selected by visually estimating the maximal leaf surface exposed to the camera.

An example of an acquisition is provided in ([Figure 4](#)), acquired in the first campaign, one day after infection. The marked leaf is one of three infected leaves in this plant and its plant stem marked with aluminum foil (appears like a black spot).



Figure 4. Thermal image from first-round experiment

3.2.2. Experiment B

A controlled experiment involving 15 grapevine plants was conducted throughout a single day in an experimental greenhouse at Evogene, Israel (31° 47' 20.472" N, 35° 12' 3.888" E) in October 2020. Images were acquired outside of the greenhouse at seven different times (rounds) between 7 a.m. and 4:30 p.m. Each round lasted about one and a half hours. In each plant, six leaves were acquired, of which three were infected with downy mildew six days after infection, and three were healthy. Each plant received two daily doses of water: one before the first round and one before the fourth round. Meteorological conditions were continuously acquired.

3.3. Equipment

Thermal images were acquired using a FLIR SC655 infrared camera (FLIR SC655, FLIR Systems, Melville, NY, USA). It uses an uncooled microbolometer detector with resolution 640×480 pixels, sensitive in the spectral range of 7.5–13 μm and possesses an accuracy of $\pm 2^\circ\text{C}$ or $\pm 2\%$ of the reading. The object temperature range is from -20 to +150 $^\circ\text{C}$.

In addition, digital images were acquired with an RGB camera (Canon EOS6D, Canon Inc., Taby, Sweden) for documentation purposes.

3.4. Databases

3.4.1. Classification of healthy and infected leaves

The full database included 1403 records (599 healthy leaves and 804 infected leaves). The records included thermographic measurements from thermal imaging of the leaves, meteorological measurements collected simultaneously, calculated features from raw data, and manual evaluation of disease severity.

To determine the earliest day that a model can detect a disease, a subset was created from the full database, which included records with actual disease severity of 5 or higher. This set included 1097 records (599 healthy leaves and 498 infected leaves). Outliers were removed from the new set, which resulted in 1012 records (571 healthy leaves and 441 infected leaves). The three data sets (the full dataset, severity of 5 or higher and after outliers' removal) are available in [Appendix D](#).

Examples of thermal and RGB images acquired at different days of infection are depicted in [Figure 5](#). Compared with the digital images, the changes in the color of the typical thermal images were more noticeable compared to the visual observations.

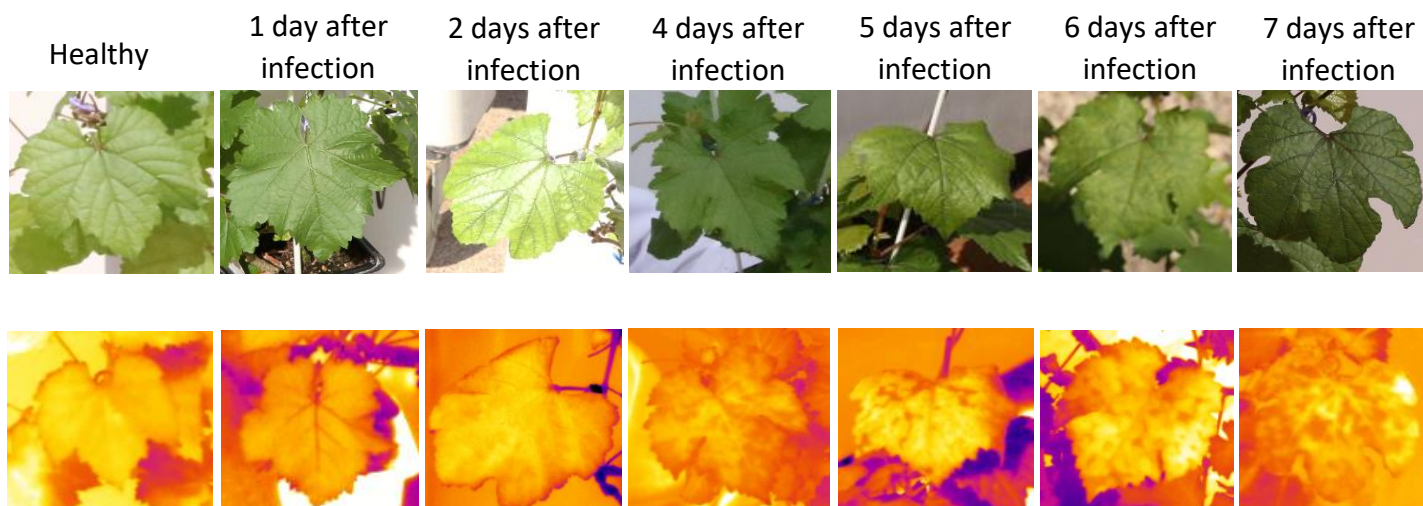


Figure 5. Examples of RGB images (top row) and thermal images (bottom row) of leaves from different infected days

3.4.2. Best acquisition time

The database contained 575 records (280 healthy leaves and 295 infected leaves). Most of the rounds included about 87 samples ([Table 2](#)). The data set is available in [Appendix E](#).

Examples of thermal images from infected and healthy leaves acquired at different times of the day are depicted in [Figure 6](#). The displayed images have the same scale (23.6°C - 34.6°C).

Round number	Acquisition Time	Number of samples
1	7:15-8:25	88
2	9:00-9:45	88
3	10:40-11:30	88
4	12:25-13:15	87
5	14:15-15:05	86
6	15:20-16:00	86
7	16:00-16:30	52

Table 2. Hours of rounds and number of samples in each round

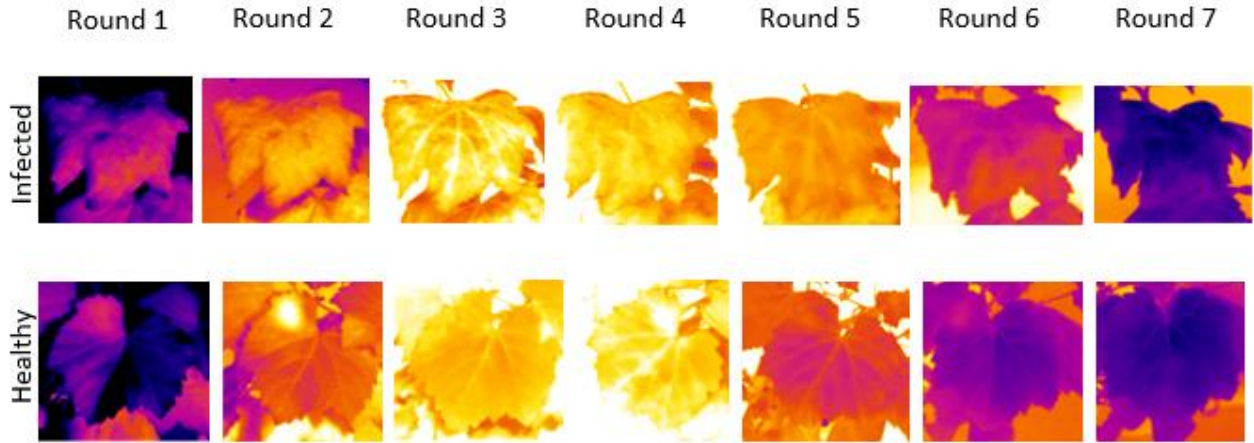


Figure 6. Thermal images from infected and healthy leaves acquired at different times of the day

3.5. Segmentation algorithm

Leaves were segmented using edge detection with the active contour algorithm (also denoted as the snakes algorithm, Kass, Witkin, & Terzopoulos, 1988). This algorithm is influenced by image forces that pull it toward features such as lines and edges. The concept of the active contour is to deform splines to the object contour in the image by minimizing an energy functional. They lock onto nearby edges and localize them accurately. This energy is classically composed of two terms, an internal and external energy term. The purpose of the internal energy term is to control the deformations made to the snake considering the regularity and smoothness of the desired contour. The purpose of the external energy term is to control the fitting of the contour onto the image, based on the intensity gradient (Cerutti et al., 2013).

The 'Chan-Vese' active contour algorithm was used (Chan & Vese, 2001). This method ensures an unbiased contour enabling to either shrink or expand based on the image features. The inputs for the algorithm were a thermal image and an initial mask; the outputs were an image with the leaf contour and a final mask.

3.6. Features

The features were calculated from the leaf mask. A description with a detailed explanation of the variables can be found in [Table 3](#). All features expressing temperature were normalized by $T - T_{air}$.

Variable Name	Description	Type	Range	Symbol	Calculation
Minimum temperature	The minimum temperature in the leaf, minus the air temperature measured at the same time	Continuous	-6.3-10.7	Tmin	Tmin-Tair
Maximum temperature	The maximum temperature in the leaf, minus the air temperature measured at the same time	Continuous	-4.2-14.6	Tmax	Tmax-Tair
Average temperature	The average of the leaf temperatures values, minus the air temperature measured at the same time	Continuous	-5.11-12.98	Tavg	Tavg-Tair
Median temperature	The median of the leaf temperatures values, minus the air temperature measured at the same time	Continuous	-5.06-13.11	median	median-Tair
Maximum Temperature Difference	The difference between the maximum and minimum temperature in the leaf	Continuous	0.5-7.1	MTD	Tmax-Tmin
Standard deviation	The standard deviation value of the leaf temperature values	Continuous	0.1-1.73	STD	std
Interquartile range	Interquartile Range (IQR), also called the midspread, is a measure of statistical dispersion and equal to the difference between 75th and 25th percentiles	Continuous	0.17-3.28	IQR	T0.75- T0.25
Median Absolute Deviation	A robust measure of the variability. Defined as the median of the absolute deviations from the median of the data	Continuous	0.1-1.53	MAD	median $\{ T_i - T_{0.5} \}$
Coefficient of variation	A standardized measure of the dispersion of a probability distribution or frequency distribution.	Continuous	0.004-0.061	Cv	STD/ Tavg
Percentile 10	The percentile is a score at or below which a given percentage fall, minus the air temperature measured	Continuous	-5.9-11.9	perc10	T0.1-Tair
Percentile 90		Continuous	-4.8-13.9	perc90	T0.9-Tair
Crop water stress index 2	A means of irrigation scheduling and crop water stress quantification based on leaf temperature measurements and prevailing meteorological conditions (Jones, 1999)	Continuous	0.37-1.53	CWSI2	$\frac{T_l - T_{wet}}{T_{dry} - T_{wet}}$
Crop water stress index 4		Continuous	0.51-30.56	CWSI4	$\frac{T_{dry} - T_l}{T_l - T_{wet}}$

Table 3. Feature description

3.7. Analysis

First, histograms and box-plots were plotted for each variable to represent the distribution and depict the groups graphically (displot and catplot, seaborn, Python). Second, outliers were removed (Python). In this research, an outlier was defined as a data point that is 1.5 times the interquartile range above the upper quartile and below the lower quartile ($Q1 - 1.5 \times IQR$ or $Q3 + 1.5 \times IQR$). Third, the correlation between the predictors was examined by Pearson's correlation coefficient (corr, pandas, Python). To avoid misleading information, the partial correlation coefficient between the predictors was also calculated (pcorr, pingouin, Python). Using a correlation coefficient to determine whether there is a numerical relationship between two variables of interest will produce misleading results if there is another, confounding variable numerically related to them both. Therefore, the partial correlation coefficient controlling the confounding variable was used (Frank, 2000). The correlation between the predictors to the response variable was examined by point-biserial correlation (biserial.cor, ltm, RStudio).

To create the supervised learning model, the significance of the variables on the response variable was evaluated by stepwise logistic regression (step, stats, RStudio). The stepwise method was "both"- a combination between "forward" and "backward" regression. Since there is no common method for feature selection, the features selected by the stepwise logistic regression were used in all models. Pearson's correlation coefficient examined the correlation between the derived features.

To give equal importance to each feature and improve the model's accuracy, quality, and learning rate, the data were normalized using the Z-score (Dhaka et al., 2021). A zero centering of data was performed by subtracting the mean value from each attribute value, then dividing each dimension by its standard deviation. A Min-Max normalization was also used, but it did not significantly improve the results ([Appendix F](#)). Normalization and standardization were conducted with MATLAB.

In order to evaluate the statistical validity of the best model, statistical tests were conducted between the best model and the second-best model. An F-test was used to test the null hypothesis that the variances of both models are equal. A T-test was used to test the null hypothesis that both models have equal means.

A sensitivity analysis was conducted for different distributions of data to explain the different classification results between the days after infection.

3.8.Classification models

Different forms of supervised learning were selected so the algorithm learns to map the inputs to the outputs by using the available input-output pairs. In other words, the aim was to ensure the mapping function should be able to estimate well the values of new input data to predict output values. Five machine learning (ML) algorithms were trained to classify infected and healthy leaves: logistic regression (LR), support vector machine (SVM), decision tree (DT), Ensemble, and Naive Bayes (NB). MathWorks MATLAB version R2019b (MathWorks Inc., Natick, MA, USA) was used to implement the ML algorithms.

- Decision tree- one of the most widely used and practical methods for inference and classification. It has a fast prediction speed and is easy to interpret. This information gain method does not assume any statistical properties of the data itself (e.g., normal distribution) and as such, it is best suited to this case where the statistical distribution is unknown. When building a decision tree, overfitting may arise which is represented in the decision tree as a deep tree with many levels. To avoid over-fitting the maximum number of splits was limited.
- Logistic regression- a statistical model that uses a logistic function to model a binary dependent variable; this was suitable in this case where there are two classes (healthy and infected).
- Naive bayes - a statistical classification technique based on Bayes Theorem. A simple supervised learning algorithm which provides a fast and accurate classification. The classifier assumes that the effect of a particular feature in a class is independent of other features. However, the algorithm still appears to work well when the independence assumption is not valid.
- Support vector machine- robust prediction models with very high accuracy of disease detection. An SVM training algorithm builds a model that assigns new examples to one category, making it a non-probabilistic binary linear classifier.
- Ensemble- the technique combines predictions from multiple machine-learning algorithms. In this work, the decision tree ensemble algorithm using the Boosting method was used. Boosting refers to a group of algorithms that trains weak learners sequentially, each trying to correct its predecessor.

3.9. Performance measures

The classification performance was quantified using the accuracy, precision, recall, F1 score and AUC (Area under the ROC Curve).

The class set contained two labels: positive (infected) and negative (healthy). Given a classifier and an instance, the four possible outcomes are ([Table 4](#)):

- True positive (TP): the leaf is infected, and it is classified as infected.
- False-negative (FN): the leaf is infected, and it is classified as healthy.
- True negative (TN): the leaf is healthy, and it is classified as healthy.
- False-positive (FP): the leaf is healthy, and it is classified as infected.

		Predicted outcome	
		Infected	Healthy
Actual value	Infected	True positive (TP)	False negative (FN) Error type 2
	Healthy	False positive (FP) Error type 1	True negative (TN)

Table 4. Confusion matrix

The accuracy determines the number of correct predictions from all the predictions. While the higher the number the better in the case of an approximately equal number of samples in all classes, accuracy alone often leads to an error in the classification of the minor class in imbalanced datasets. Accuracy is computed as:

$$Accuracy = \frac{TP + TN}{Total\ number\ of\ instances} \quad (1)$$

The precision is the positive predictive value and is computed as:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

The recall is also referred to as the true positive (TPR) rate and is computed as:

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

The F1 score is the harmonic mean of the precision and recall and is computed as:

$$F1\ score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

A ROC curve (receiver operating characteristic curve) is a graph showing the performance of a classification model at all classification thresholds displaying two parameters:

- Recall (also TPR)
- False Positive Rate (FPR)

False Positive Rate is defined as follows:

$$FPR = \frac{FP}{FP + TN} \quad (5)$$

A ROC curve plots TPR vs. FPR at different classification thresholds. Lowering the classification threshold classifies more items as positive, thus increasing both False Positives and True Positives.

AUC stands for "Area under the ROC Curve". The AUC measures the entire two-dimensional area underneath the entire ROC curve.

4. Algorithms

4.1. Edge detection

The 'Chan-Vese' active contour algorithm was applied for leaf detection. The software was implemented using MATLAB ([Appendix G](#)) version R2019b (MathWorks Inc., Natick, MA, USA) with additional functions (Shawn Lankton, 2007).

First, the thermal image was converted to grayscale. Second, the range of pixel values was changed by cutting the histogram edges ([Figure 7](#)) to clear the color contrast. This change was made for visualization purposes only.

Third, the active contour algorithm was activated on the image. The process is depicted ([Figure 8](#)). Initial mask position and size were manually set for each leaf. Image and initial mask were the algorithm inputs (Fig. 8 A(a), 8 A(b)). The algorithm performed iterations to find the contours of the leaf (Fig. 8 A(c)) with a Lambda parameter of 0.3. Lambda is the weight of the smoothing term (a higher value reaches a smoother term). After a maximum of 200 iterations ([Appendix H](#)) a final mask was obtained (Fig. 8 A(d)). The output was an image of the leaf contour (final mask, Fig. 8 B). The returned final mask was a matrix where pixels inside the mask were set to leaf temperature and pixels outside the mask were set to zero. Visualizations of all images after the active contour algorithm are displayed in [Appendix I](#).

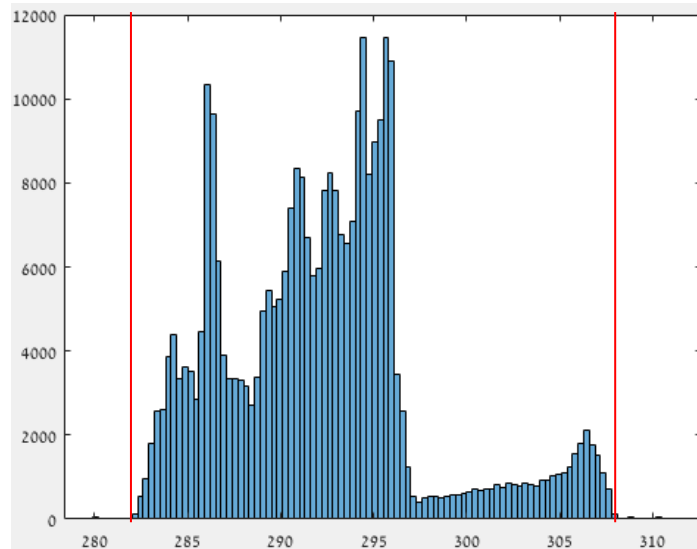


Figure 7. Prevalence of pixels of the entire image

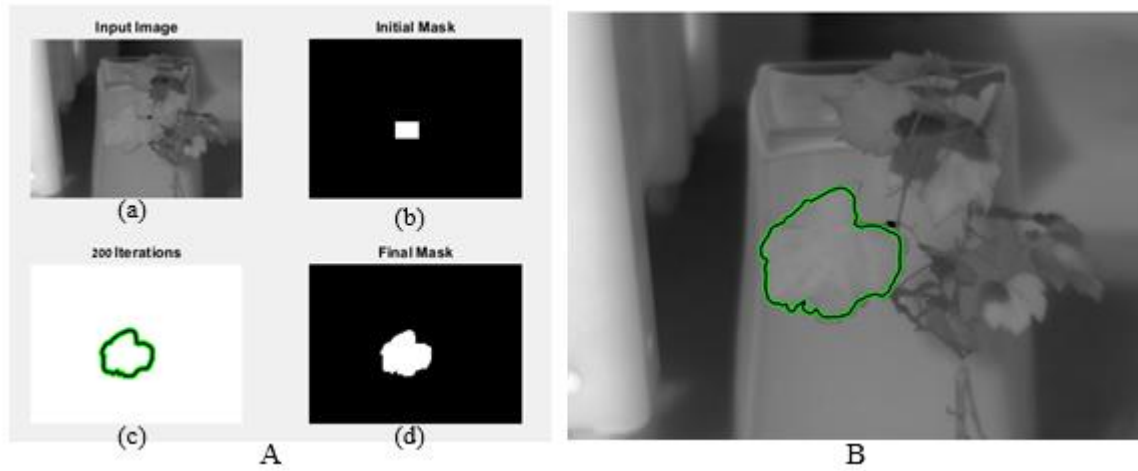


Figure 8. The output of the active contour algorithm. Includes image, initial and final mask (A) and image with final mask (B)

4.2. Classification models

Classification was performed by K-fold cross-validation (K=5) for each of the five models. The parameters for each model were optimized using MATLAB version R2019b ([Appendix J](#)). The optimal parameters obtained for each model were:

- Decision Tree: 17 maximum splits, Maximum deviance reduction split criterion
- Logistic Regression: Threshold 0.5
- Naive Bayes: Gaussian kernel
- SVM: Cubic Kernel, box constraint level of 1
- Ensemble: LogitBoost method, 14 maximum splits, a learning rate of 0.348

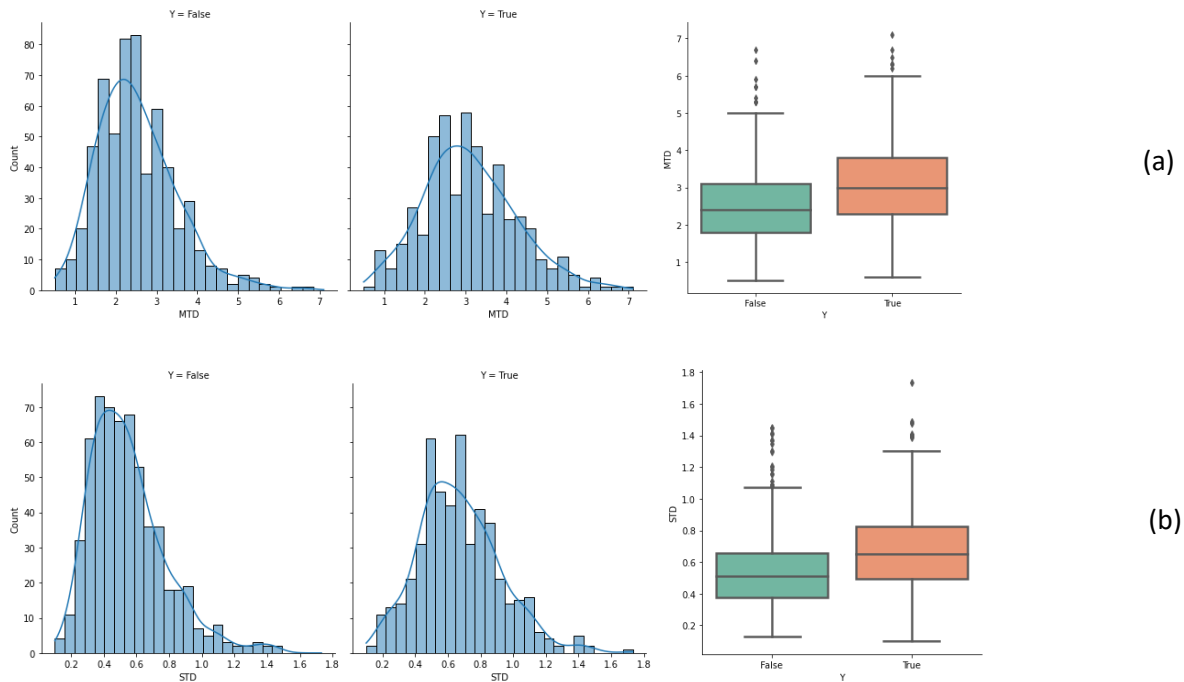
5. Results and Discussion

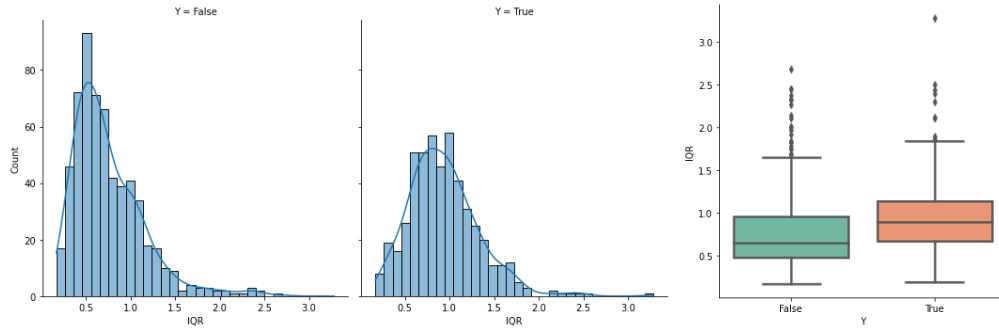
5.1. Classification of healthy and infected leaves

5.1.1. Histograms and boxplots

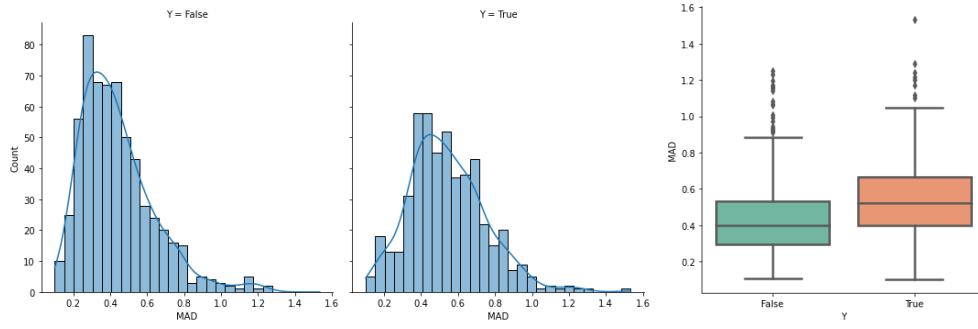
Histograms and boxplots of healthy (false) and infected leaves (true) were created for each variable for data visualization (Software code is detailed in [Appendix K](#)). [Figure 9](#) presents the following variables: MTD (Fig. 9a), STD (Fig. 9b), IQR (Fig. 9c), MAD (Fig. 9d), Tavg (Fig. 9e), Tmin (Fig. 9f), Tmax (Fig. 9g), median (Fig. 9h), perc10 (Fig. 9i), perc90 (Fig. 9j), CWSI2 (Fig. 9k), CWSI4 (Fig. 9l), and Cv (Fig. 9m). The boxplots below indicate that the predictors Tavg, Tmin, Tmax, median, perc10, and perc90 are different for infected and healthy leaves as their median line of the box-plot lies outside of the box of comparison box plot (higher for infected leaves). The histograms reveal that most predictors follow a normal distribution. Also, the histograms reveal that there is more scatter in the infected leaves and with a stronger mean signal.

Feature CWSI was calculated in two ways, representing the same variable. The histogram of CWSI4 is right-skewed with a long tail. Therefore, it was decided to select CWSI2 for the feature selection model.

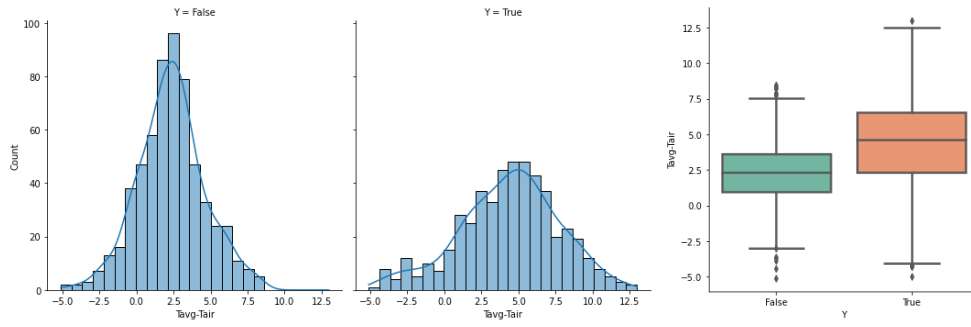




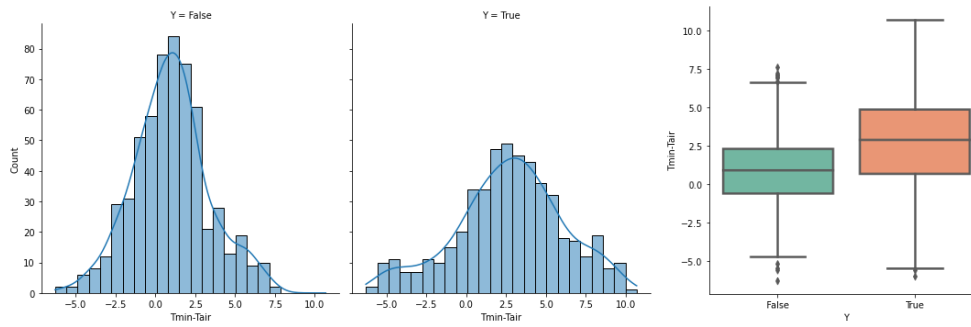
(c)



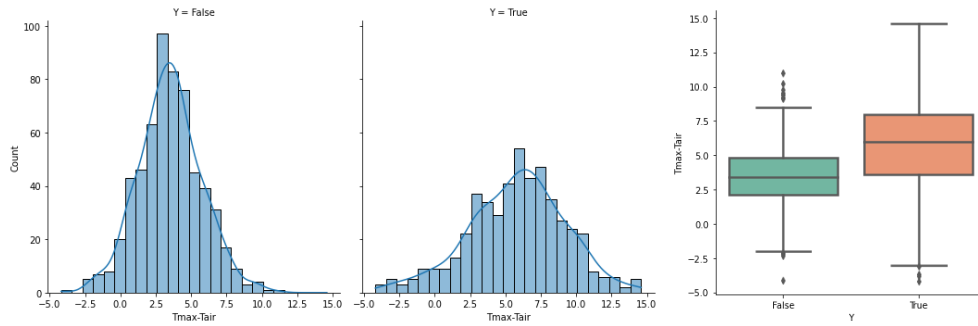
(d)



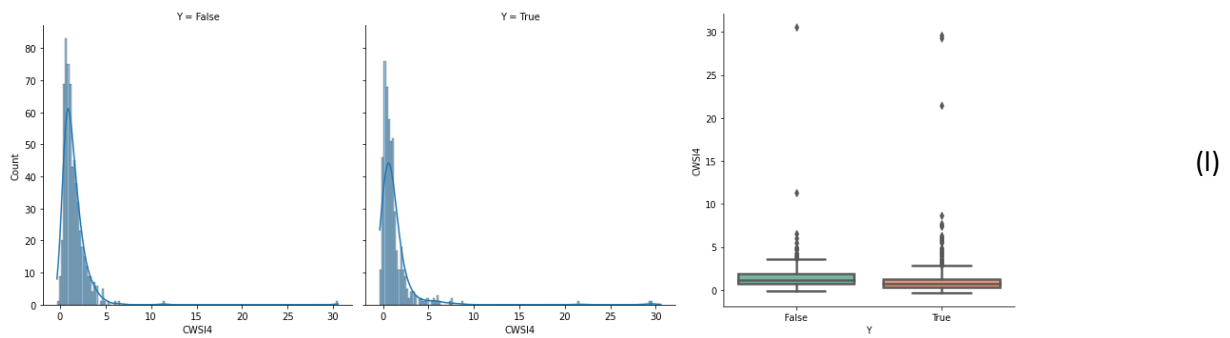
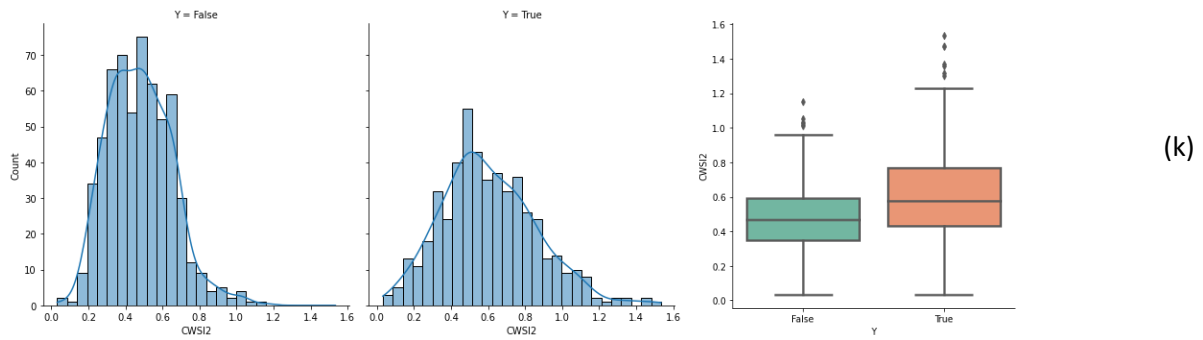
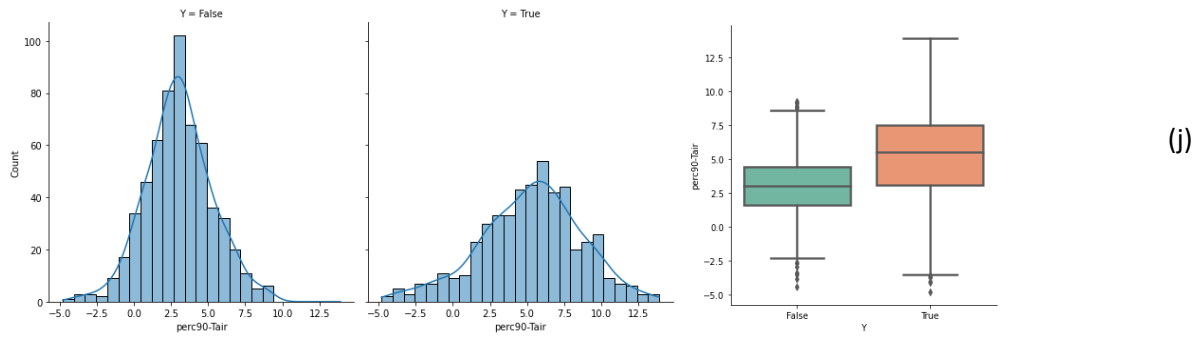
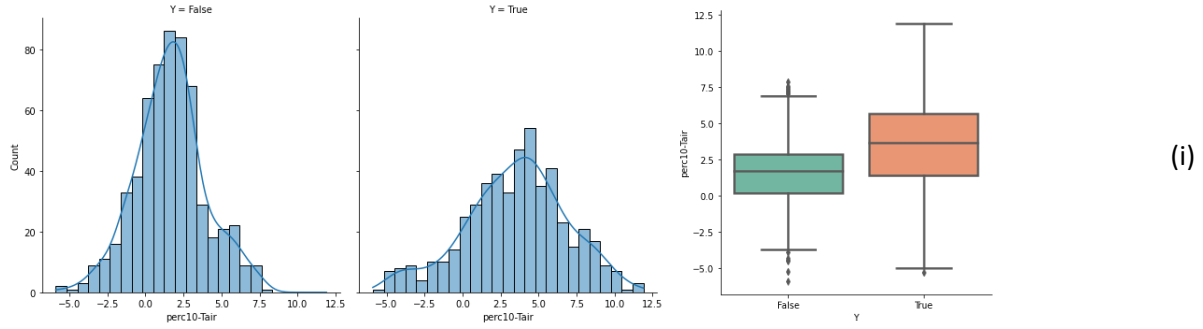
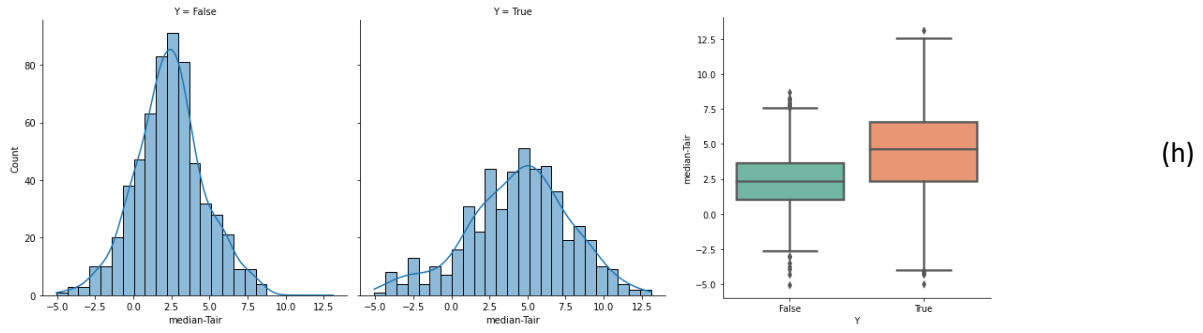
(e)



(f)



(g)



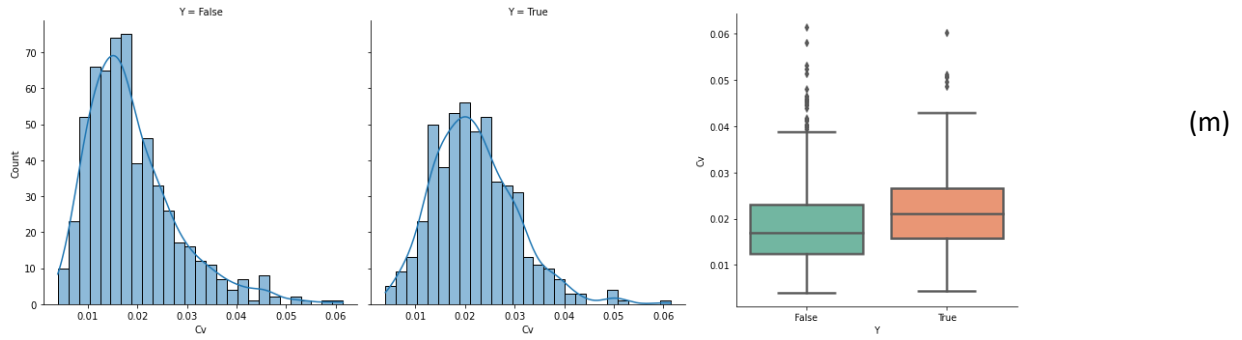


Figure 9. Histograms (left) and boxplots (right) for each variable (a-m)

5.1.2. Outliers

When reviewing the box-plots of all features ([Figure 10](#)) of 1097 records, it can be seen that there are several data points located outside the whiskers of the box plot.

After removing the outliers that were 1.5 times the interquartile range above the upper quartile and below the lower quartile (boxplot of features in [Figure 11](#)), 1012 records were left. Software code is detailed [Appendix K](#).

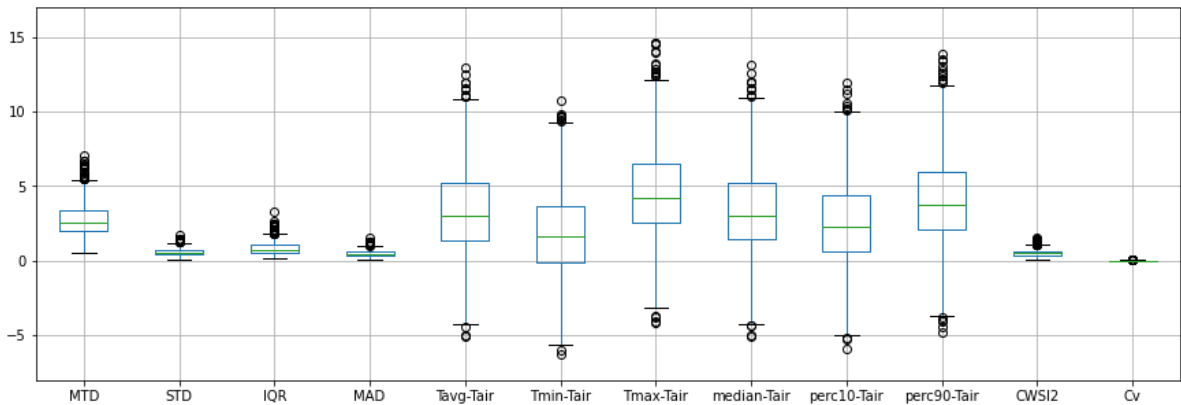


Figure 10. All features box-plots

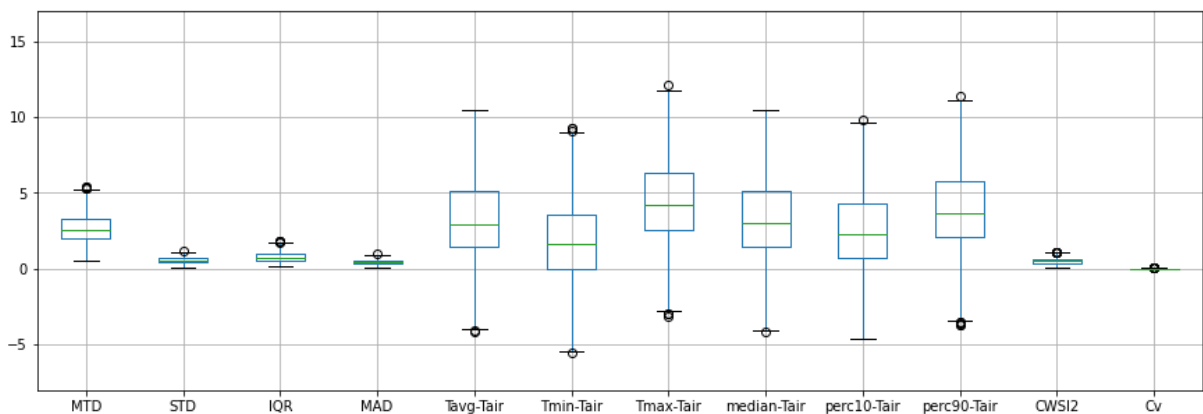


Figure 11. All features box-plots without outliers

5.1.3. Correlation analysis

5.1.3.1. Pearson correlation

The displayed Pearson correlation matrix (Figure 12) shows that some of the variables are highly correlated with each other. Thus, there is multicollinearity between the predictors. However, this result was expected due to the variable construction method. The high correlations shown in the lower right dark square originate from the fact that they are all temperature-related. The high correlations in the higher left square originate from the fact that these predictors are related to the dispersion of the data. Also, there are low correlations between these two groups.

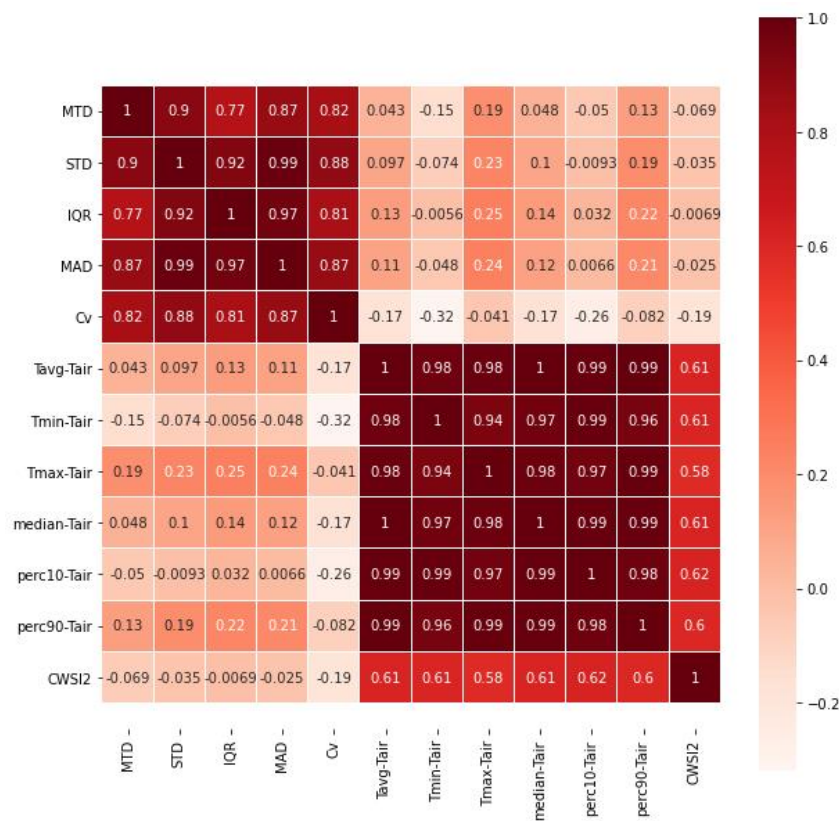


Figure 12. Heatmap of Pearson Correlation Coefficient Matrix

5.1.3.2. Pearson Partial correlation

The partial correlation (Figure 13) measures the degree of association between two random variables, with the effect of a set of controlling random variables removed. Comparing the partial correlations with the correlations presented above reveals the following differences. Firstly, the partial correlations between predictors and the dependent variable (Y) are much lower as compared to the biserial correlations. Also, when removing all other predictors

effects – MTD, perc90, and CWSI2 have the highest correlation with the outcome variable. CWSI2 was ranked in the penultimate place at biserial correlations and in the Partial correlation, it is third. The lowest partial correlations obtained with the outcome variable are perc10 and Tmin. Here, too, perc10 is rated relatively high in the biserial correlations and is now rated among the latter. Moreover, the partial correlations between the predictors are also lower than the previously obtained correlations (Pearson). Not only that the absolute size of the correlations dropped when removing the other predictors effects but that also the direction of some correlations changed. An example is given by the correlation between perc10 and perc90, which is almost perfectly correlated in the full Pearson correlation ($r = 0.98$), and in the partial correlation is much lower and changes its direction ($r_p = -0.039$).

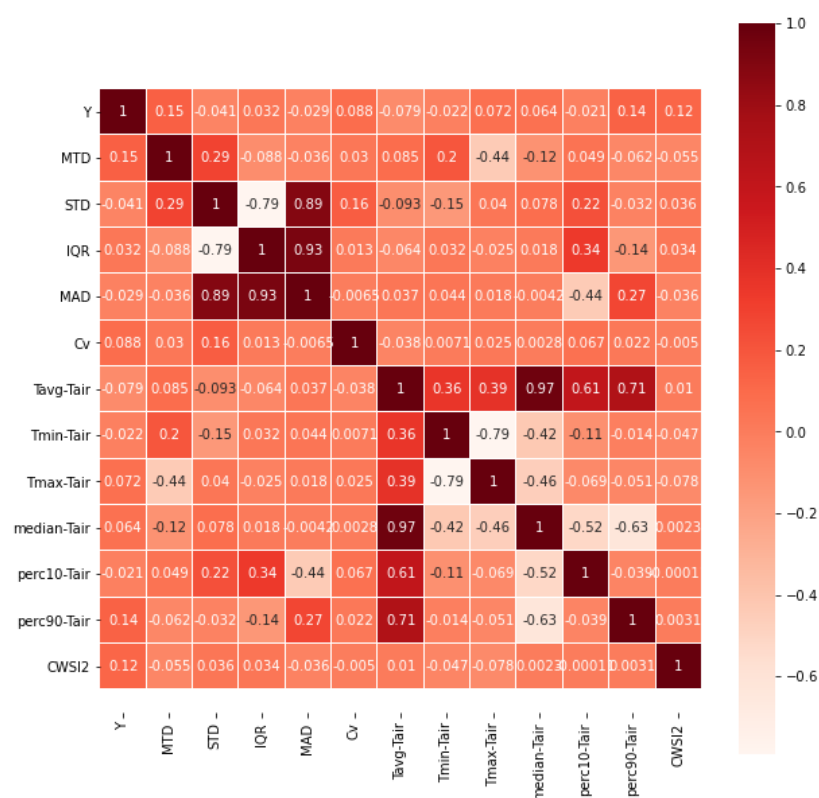


Figure 13. Pearson's partial correlations

5.1.3.3. Point-biserial correlation

The point-biserial correlation was used to check correlations between the predictors and the dependent variable (Table 5); the coefficients of correlation ranged from $r_{Cv} = 0.2086$ to $r_{Tmax} = 0.3689$, where Tmax had the highest correlation and Cv the lowest.

Variable	Biserial correlation
Tmax	0.3689
perc90	0.3546
Tavg	0.3279
Median	0.3270
perc10	0.3001
MAD	0.2887
MTD	0.2863
IQR	0.2854
STD	0.2823
Tmin	0.2746
CWSI2	0.2582
Cv	0.2086

Table 5. Point-biserial correlations

The code for determining the correlations is detailed in [Appendix K](#).

5.1.4. Model analysis

5.1.4.1. Feature selection

The feature selection conducted by stepwise logistic regression ([Appendix L](#)) resulted in seven significant variables: MTD, STD, Cv, Tavg, median, perc90, CWSI2. This result was unexpected because Pearson's partial correlation matrix showed that Tavg and median were in a high positive linear correlation ($r=0.97$), which led to the expectation that the model would not include both. Another unexpected result was the estimate of Tavg ($\beta=-9.53$) which had the opposite sign of the median ($\beta=3.696$). A possible reason for these results is that the selected method for stepwise logistic regression ("both") begins with the full model. When the two variables (Tavg and median) are together in the model and have high multicollinearity and opposite sign, they cancel each other out and stay in the model instead of removing both. Therefore, both variables were removed from the model manually. This led to a higher AIC, resulting with a better model. Finally, to train the classification models the five remained significant variables were selected: MTD, STD, Cv, perc90, CWSI2.

5.1.4.2. Classification analysis

The accuracy, precision, recall, F1 score and AUC (confusion matrix and ROC curve, [Appendix M](#)) of all constructed models ([Table 6](#)) reveal that the best results were achieved by the SVM model with major differences between the performance measures of the other models.

The results of the F-test ([Appendix N](#)) between the best model (SVM) and the second-best model (NB) using the F1 scores of each of the five folds (cross-validation) revealed that the

variances of both models were equal ($F (=3.27) < F$ Critical one-tail ($=6.39$)) implying that the null hypothesis was not rejected.

The two-tail T-test assuming equal variances ([Appendix N](#)) revealed that the null hypothesis was rejected (t Stat ($=-5.56$) $< t$ Critical two-tail ($=-2.31$)). The observed difference between the sample means ($0.77 - 0.67$) suggests a significant difference between the models. SVM was chosen to be the best model.

To determine the earliest day for downy mildew detection, the results in the following table ([Table 7](#)) obtained by the SVM were detailed according to each of the 'days after infection'.

The results revealed that later days after infection produced lower results than earlier ones, contrary to expectations (since the disease develops over time and should be easier to detect as time passes). To determine the cause of this, a series of analyses as described in [Figure 14](#), were conducted based on different explanations that may have led to these results.

Model Measure	Decision Tree	Logistic Regression	NB	SVM	Ensemble
F1 score	60.5%	64.9%	66.8%	77.5%	63.8%
Precision	70.5%	70.8%	69.7%	83.1%	70.9%
Recall	53.1%	59.9%	64.2%	71.6%	58%
AUC	0.728	0.762	0.783	0.874	0.776
Accuracy	69.9%	71.7%	72.2%	81.6%	71.3%

Table 6. All results from all models based on all data

Days after infection	Number of samples	Number of misses	Accuracy
0	571	65	88.6%
1	19	2	89.5%
2	61	3	95.1%
4	180	55	69.4%
5	39	4	89.7%
6	44	17	61.4%
7	98	40	59.2%

Table 7. Results by day after infection

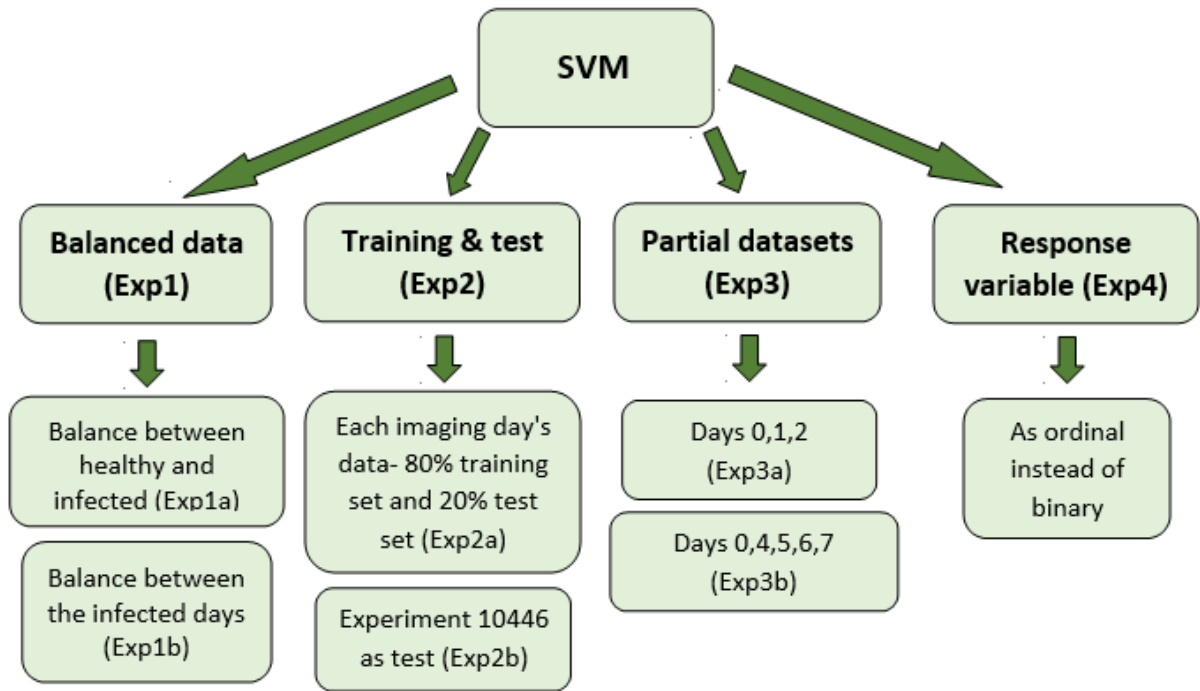


Figure 14. Diagram representing the logical flow of the different analyses

1. To avoid bias, the dataset was balanced.
 - a. By removing records from the healthy leaves, the samples number between the diseased and healthy leaves was balanced ([Table 8](#)). The accuracy achieved with these balance results was 79.1%, the F1 score was 77.9%, and the AUC was 0.86. The model's accuracy of the later days after infection improved. In order to further improve the results, a new feature selection was performed on the balanced data resulting in an accuracy of 79.2%, an F1 score of 77.8%, and an AUC of 0.86. The results were pretty similar.

Days after infection	Number of samples	Number of misses	Accuracy
0	441	68	84.6%
1	19	1	94.7%
2	61	5	91.8%
4	180	55	69.4%
5	39	5	87.2%
6	44	14	68.2%
7	98	36	63.3%
New features- MTD, SVM, perc10, perc90, Cv, CWSI2			
0	441	63	85.7%
1	19	2	89.5%
2	61	3	95.1%
4	180	56	68.9%
5	39	6	84.6%
6	44	15	65.9%
7	98	38	61.2%

Table 8. Results after balancing healthy and infected samples (top), and with this balanced sample after selecting new features (bottom) (Exp1a)

- b. Besides taking similar numbers of samples from healthy and infected leaves, the number of samples taken each day after infection was also balanced ([Table 9](#)). The accuracy achieved was 73.8%, the F1 score was 71%, and the AUC was 0.756. This did not yield improvement in the model's results, but the accuracy of the later days after infection improved greatly. The accuracy achieved after a new feature selection was 70.2%, the F1 score was 67.1%, and the AUC was 0.734. The previous features performed better.

Days after infection	Number of samples	Number of misses	Accuracy
0	84	14	83.3%
1	14	7	50%
2	14	4	71.4%
4	14	10	28.6%
5	14	3	78.6%
6	14	4	71.4%
7	14	2	85.7%
New features- MAD, perc90, Cv, CWSI2			
0	84	17	79.8%
1	14	6	57.1%
2	14	4	71.4%
4	14	11	21.4%
5	14	5	64.3%
6	14	7	50%
7	14	0	100%

Table 9. Results after balancing between the infected days (top), and with this balanced data after selecting new features (bottom) (Exp1b)

2. To examine the effects of different climatic conditions on the results, the data was divided into different training and test sets.
 - a. As the climatic conditions differed between imaging days, it was difficult to classify the data. Each imaging day's data was split in two: 80% from the data for the training set and 20% from the data for the test set ([Table 10](#)). The accuracy of the training set was 82.5%, with an F1 score of 78.3% and an AUC of 0.886. The test set accuracy was 76.5%, with an F1 score of 70.8% and an AUC of 0.827. Some days' results improved, while others did not. The model's performance did not improve.

Days after infection	Number of samples	Number of misses	Accuracy
0	112	16	85.7%
1	5	1	80%
2	12	0	100%
4	36	16	55.6%
5	8	0	100%
6	8	6	25%
7	19	8	57.9%

Table 10. Results of the test set (Exp2a)

- b. The test set included one specific experiment (No. 10446), and the training set included the rest. Experiment 10446 included the days 0,4,5,6,7 after infection. Days 5,6 were not included in any other experiments, so they appear only in the test set ([Table 11](#)). The accuracy of the training set was 86.4%, the F1 score was 82% and the AUC was 0.892. The test set accuracy was 57.8%, with an F1 score of 48.9% and an AUC of 0.593. The training and test sets were very different in their results. Even on the days that appear in the training set (0,4,7), the results are poor. According to this analysis, it is not possible to classify healthy and downy mildew infected from images acquired on an imaging day that is not included in the training set.

Days after infection	Number of samples	Number of misses	Accuracy
0	171	43	74.8%
4	43	24	44.2%
5	39	19	51.3%
6	44	28	36.4%
7	44	30	31.8%

Table 11. Results of the test set (experiment No. 10446) (Exp2b)

3. Considering that the accuracy on days 1 and 2 after infection was greater than that of the other days, it was tested whether these days affect the prediction.
- a. The following hypothesis was tested: whether with a dataset of leaves from healthy and 1 and 2 days after infection, the accuracy of the prediction for these days still remains high ([Table 12](#)). The data for this analysis was balanced. The accuracy achieved was 91.9%, the F1 score was 92.1%, and the AUC was 0.961. The results showed that even without the later days, the early days' predictions held true very well. In order to further improve the results, a new feature selection was performed on the data. The accuracy achieved was 92.9%, the F1 score was 92.8%, and the AUC was 0.969. The new features improved the results.

Days after infection	Number of samples	Number of misses	Accuracy
0	99	10	89.9%
1	21	1	95.2%
2	78	5	93.6%
New features- STD, median, perc10, perc90, Cv, CWSI2			
0	99	5	94.9%
1	21	4	80.9%
2	78	5	93.6%

Table 12. Results of a dataset with days 0,1,2 after infection (top), and the results of this same dataset with new features (bottom) (Exp3a)

- b. The same was done by using a dataset of days 0,4,5,6,7 after infection ([Table 13](#)). Here also the data was balanced. The accuracy achieved was 79.1%, the F1 score was 78.4%, and the AUC was 0.856. The results showed that when days 1 and 2 after infection were removed, the classification results of the later days improved. The accuracy achieved after a new feature selection was 76.7%, the F1 score was 75.7%, and the AUC was 0.833. The previous features performed better.

Days after infection	Number of samples	Number of misses	Accuracy
0	399	72	81.9%
4	197	54	72.6%
5	45	5	88.9%
6	45	9	80%
7	112	27	75.9%
New features- MTD, median, perc10, perc90, Cv			
0	399	77	80.7%
4	197	72	63.4%
5	45	6	86.7%
6	45	11	75.5%
7	112	20	82.1%

Table 13. Results of a dataset with days 0,4,5,6,7 after infection (top), and the results of this same dataset with new features (bottom) (Exp3b)

4. Since the results of the other approaches varied each day after infection, an analysis was performed to examine if the response variable should be ordinal instead of binary. The assumption was that disease development increased every day, suggesting some kind of order as imaging days progressed. Therefore, an ordinal regression was

performed on the imbalanced data. A new feature selection was conducted to accommodate for the new response variable and resulted in the following features: MTD, IQR, MAD, median, perc10, perc90, and CV. The accuracy of the ordinal classification was 65.4% with many observations classified as healthy even though they were infected (and were not classified as infected on another day, [Figure 15](#)).

Since it does not matter which day after infection was classified, but whether the leaf was healthy or infected, the results were converted to binary so all days that were not classified as 0 (healthy) were deemed infected ([Table 14](#)). This classification of the converted response variable resulted in an accuracy of 74.9%. The binary response variable performed better than both the ordinal response variable and the converted response variable and hence was selected.

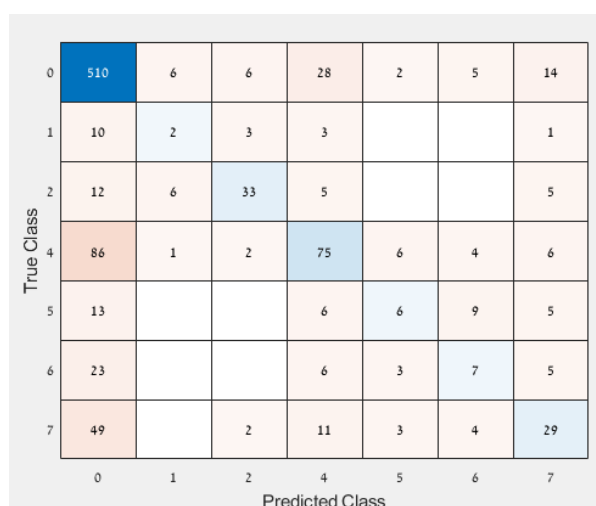


Figure 15. Confusion matrix of ordinal regression

Days after infection	Number of samples	Number of misses	Accuracy
0	571	61	89.3%
1	19	10	47.4%
2	61	12	80.3%
4	180	86	52.2%
5	39	13	66.7%
6	44	23	47.7%
7	98	49	50%

Table 14. Results of ordinary regression (Exp4)

5.1.5. Misclassifications

In the different datasets, some images were repeatedly classified incorrectly. For each infection day, one example of correct and incorrect classification is displayed ([Figure 16](#)). There is no apparent reason for the differences in the classification.

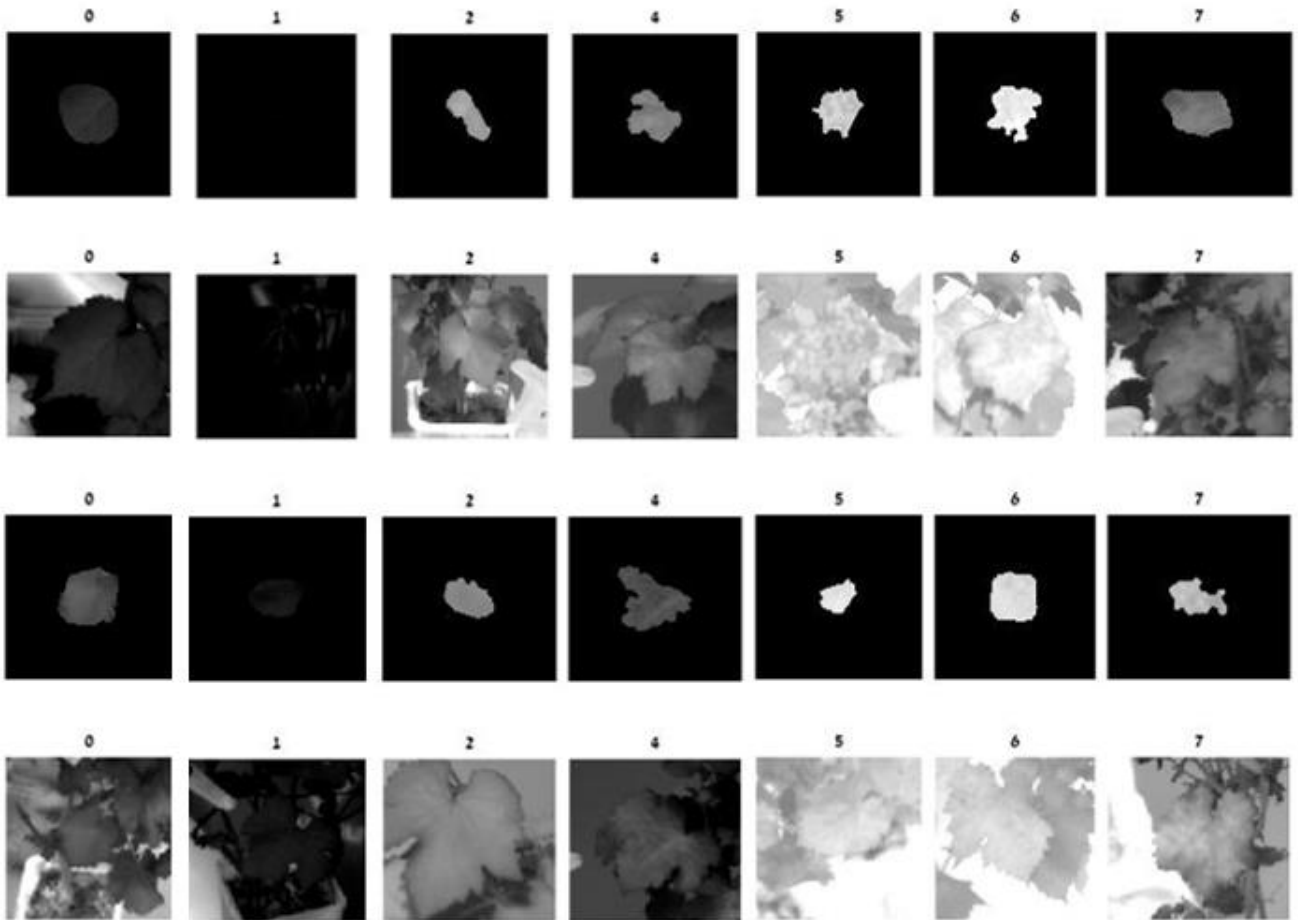


Figure 16. The first two rows of images have been correctly classified. The first row shows the masks of the Active Contour algorithm, resulting from the image in the second row. The same is true for rows three and four, but these are examples of incorrect classification.

5.1.6. Summary

Out of the five developed models the SVM model produced best results. Seven different approaches were used to explain the large differences in disease detection between days after infection ([Table 15](#)). Although the best results were with an SVM model using a dataset with healthy and infected leaves from only 1 and 2 days after infection, this dataset itself is not selected since it included only the two first days after infection without the remaining days. These good results were probably due to the fact that the plants were disturbed by the controlled infection which was what was probably actually detected.

Additionally, in each approach, the new selected features were very similar to the first selection, with performance differing by only $\pm 3\%$. These results indicated that the first feature selection was indeed suitable for the different datasets.

The best relevant results were obtained from a SVM model using a balanced dataset of healthy leaves and infected leaves.

A possible explanation for the differences in disease detection between days after infection might be that infrared thermography measurements are influenced by many other factors that affect leaf temperature, including ambient temperature, humidity, sunlight, and wind (Granum et al., 2015). The images were acquired at different dates, times of day, and under different environmental conditions, which may have affected the characteristics of the images or even the plants. Although each variable related to temperature has been normalized (T-Tair), this might not be sufficient to accommodate environmental changes.

The classification results of days one and two after infection, which were better than those of the other days, was probably not because the downy mildew was detected early; It is more likely that the infection itself, which was manually induced, may have affected the leaves locally and caused a strong physiological reaction to the leaves. As the virus penetrated the plant, the local effect diminished.

Despite many different approaches and analyses conducted, we could not provide an explanation for the inconsistent results between the days (days 4-7 after infection).

Since each of the models (the model with data of days 0,1,2 and the model with data of days 0,4,5,6,7) provided good results, future research should consider dividing the data into 3 groups: healthy, infected at first days, and infected at late days.

Approach	F1 score	AUC	Accuracy	New features F1 score	New features AUC	New features accuracy
SVM- all data	77.5%	0.874	81.6%			
Balance between healthy and infected (Exp1a)	77.9%	0.86	79.1%	77.8%	0.86	79.2%
Balance between the infected days (Exp1b)	71%	0.756	73.8%	67.1%	0.734	70.2%
Each imaging day's data- 80% training set and 20% test set (Exp2a)	70.8%	0.827	76.5%			
Experiment 10446 as test (Exp2b)	48.9%	0.593	57.8%			
Days 0,1,2 (Exp3a)	92.1%	0.961	91.9%	92.8%	0.969	92.9%
Days 0,4,5,6,7 (Exp3b)	78.4%	0.856	79.1%	75.7%	0.833	76.7%
As ordinal instead of binary (Exp4)			74.9%			

Table 15. A summary of all approaches and their results

5.2. Best acquisition time

5.2.1. Histograms

Histograms were plotted for each variable revealing differences between the different sampling times along the day ([Figure 17](#)). The histograms are numbered from round number 1 (left, 1.5 hours after dawn) to round number 7 (right, late afternoon).

5.2.2. Model analysis

[Table 16](#) presents accuracy, F1 score and AUC obtained by SVM model for all rounds. Best results were obtained when data were collected 10:40-11:30 (round 3) with accuracy differences of about 6% and 23% from the next best and worst rounds respectively. These results are different from results of Alchanatis et al. (2010), who found that midday (12:00-14:00) was the best time to map and estimate water status variability using thermal imaging. Although these are two different problems, they are similar in expectation to see temperature differences within the leaves.

The results of the F-test ([Appendix O](#)) between the best round (3) and the second-best round (1) using the accuracy of each of the five folds from cross-validation (since the data

was balanced) revealed that the null hypothesis was not rejected ($F (=4.08) < F$ Critical one-tail ($=6.39$)), implying that the variances of both rounds were equal.

The two-tail T-test assuming equal variances ([Appendix O](#)) revealed that the null hypothesis was not rejected ($-t$ Critical two-tail ($=-2.31$) $< t$ Stat ($=-1.12$) $< t$ Critical two-tail ($=2.31$)). The observed difference between the sample means ($0.807 - 0.751$) is not convincing enough to say that the results differ significantly.

Although the 3rd round was not statistically better, records from these hours were analyzed to generalize the findings since best results were obtained at this acquisition time. A special new dataset of 239 records (95 healthy and 144 infected) was created using only observations that were acquired 10:40 and 11:30 a.m. from the classification model (dataset described in [section 3.4.1](#)). This dataset was used to create a new model; the results are detailed in [Table 17](#). To improve the results, a new feature selection was performed. The selected features were: MTD, Tavg, Tmin, perc10, perc90, and Cv. The new features improved the results and the model produced good results ([Table 17](#)).

As compared with the model that used all the data, the model from the best hours included fewer observations, but the results were quite similar. This model would likely perform even better if it were trained on a larger dataset. Therefore, it is worth considering acquiring images at similar times along the day, even if they are acquired on different days. Further research is needed on this topic.

Though, even when the acquisition occurred during the best hours, results are strongly affected by the environmental conditions, which may change day by day, and impact the results.

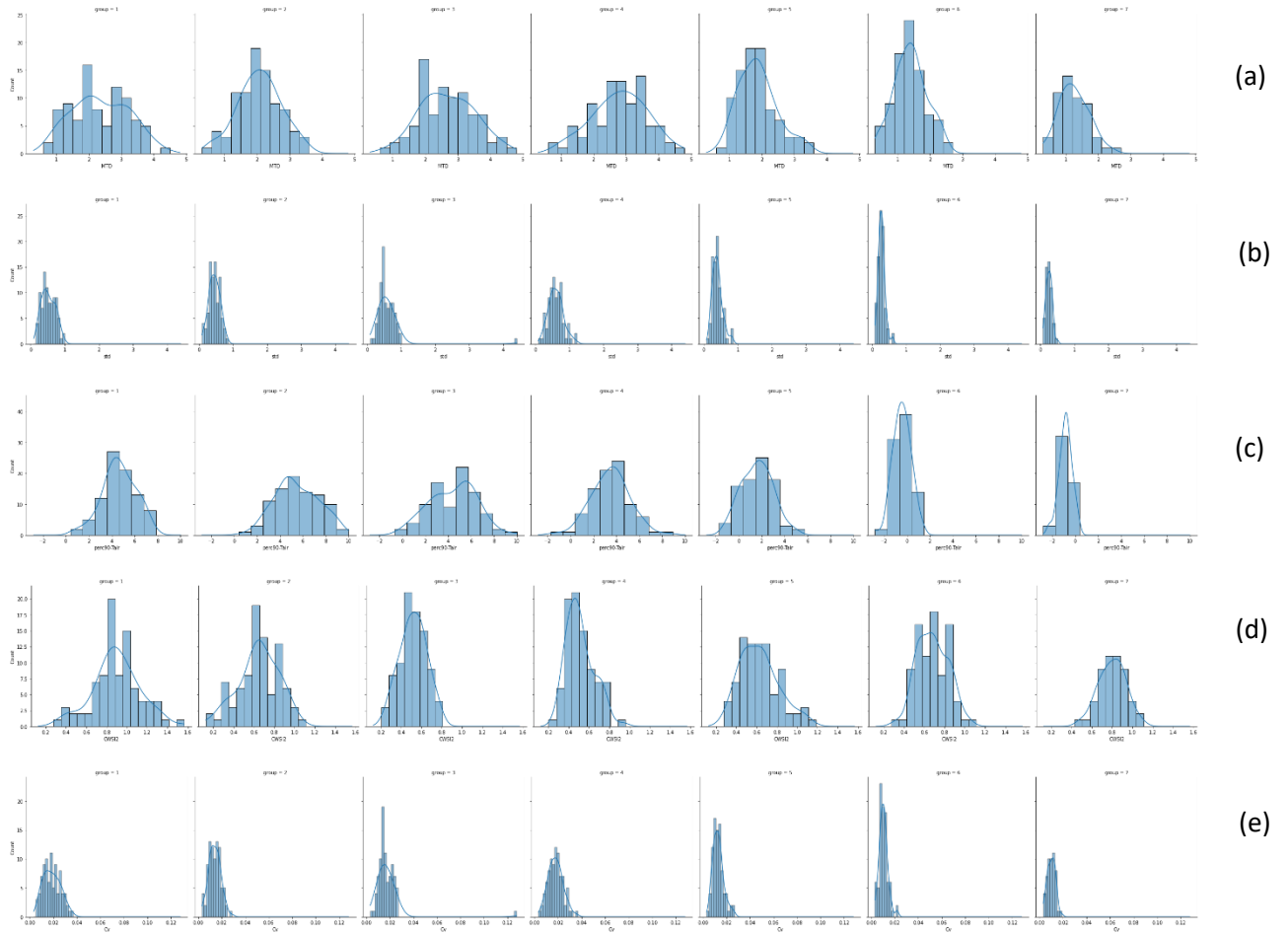


Figure 17. Histograms of each selected variable:
(a) MTD, (b) STD, (c) perc90, (d) CWSI2, (e) Cv

Measure Round No./Time	Accuracy	F1 score	AUC
(1) 7:15-8:25	75%	75.6%	0.774
(2) 9:00-9:45	72.7%	72.7%	0.794
(3) 10:40-11:30	80.7%	80.5%	0.895
(4) 12:25-13:15	59.8%	61.5%	0.676
(5) 14:15-15:05	65.1%	67.4%	0.691
(6) 15:20-16:00	58.1%	61.7%	0.644
(7) 16:00-16:30	57.7%	59.3%	0.557

Table 16. Performance for each round of the diurnal measurements

Model	Number of samples	F1 score	AUC	Accuracy
Hours 10:40-11:30	239	72.7%	0.764	67.4%
New features	239	80.8%	0.826	76.6%

Table 17. Results of a dataset from 10:40 to 11:30 and after a new feature selection

6. Conclusions

The best model for classifying between healthy and infected leaves was a SVM model with the following features: MTD, STD, perc90, Cv and CWSI2. The model resulted in 10% higher performance than all other models tested (81.6% accuracy, 77.5% F1 score and 0.874 AUC).

The inconsistent results between the days (days 4-7 after infection) could not be explained.

The best time of day for acquiring images for downy mildew detection was between 10:40 and 11:30 resulting in 80.7% accuracy, 80.5% F1 score, and 0.895 AUC.

The results indicate that thermograms can detect downy mildew infection, even before any visible symptoms appear. Using images from the best hours can improve performance, even if the images are not from the same days.

However, even when the image acquisition is conducted at the best time, variations in illumination cannot be avoided, which results in reduced performance. There is a trade-off between using a large and wide database (acquired along many dates) and detecting the disease (the fewer dates, the easier).

Appendices

Appendix A – Journal submitted paper

Early detection of grapevine downy mildew and diurnal variations using thermal imaging

Bar Cohen^{1,2}, Yael Edan², Asher Levi¹, Victor Alchanatis¹

Abstract

The agricultural industry is facing a serious threat from plant diseases that cause production losses and economic losses. Early information on disease development can improve disease control using suitable management strategies. This study sought to detect downy mildew on grapevine leaves at early stages of development using thermal imaging technology and to determine the best time along the day for image acquisition. In controlled experiments, 1587 thermal images of grapevines grown in a greenhouse were acquired before infection, 1, 2, 4, 5, 6, and 7 days after a controlled infection, around midday. Also, images of healthy and infected leaves were acquired at seven different times along the day between 7 a.m. and 4:30 p.m.

Leaves were segmented using the active contour algorithm. Twelve features were derived from leaf mask and meteorological measurements. Stepwise logistic regression revealed five significant features used in five classification models. Performance was evaluated using K-folds cross-validation. The support vector machine model produced the best classification accuracy of 81.6%, F1 score of 77.5% and AUC of 0.874. Acquiring images in the morning between 10:40 and 11:30 a.m., resulted in 80.7% accuracy, 80.5% F1 score, and 0.895 AUC.

Keywords: Precision agriculture, Disease detection, Classification

Introduction

Plant diseases are a major cause of production losses and economic losses in the agriculture industry (Savary *et al.*, 2012). Pathogens are responsible for direct yield losses of 20-40% of global agricultural productivity (Savary *et al.*, 2012). In order to ensure sustainable agriculture, it is essential to monitor plant health to prevent disease spread with as little damage to crop production as possible. However, the main challenge is the difficulty in determining the physical, chemical and biological changes in plants before symptoms of infection appear (Cui *et al.*, 2018).

Disease detection techniques can be classified into invasive and non-invasive methods. Invasive techniques involve destructive leaf sampling followed by chemical treatments after direct identification of the pathogen (Sankaran *et al.*, 2010). Non-invasive techniques identify plant diseases by detecting the impact of the pathogen on the physiological plant response. Currently, the most promising non-invasive disease detection methods are sensors that measure temperature, reflectance, or fluorescence (Mahlein, 2016).

Leaf temperature is a rapid response variable that can reveal crop stresses before visible symptoms appear (Khanal *et al.*, 2017). Stressed plants respond with physiologic protection mechanisms that lead to changes in leaf surface temperature (Mahlein, 2016). Infrared thermography (IRT) enables the evaluation of the plant temperature related to changes in water status and transpiration due to infections by pathogens. Temperature differences within individual leaves, plants, and crops indicate the presence of disease in plants (Mahlein, 2016). Compared to optical, multispectral, and hyperspectral sensors, thermal sensors have shown to be more effective at detecting disease-induced early modifications (Khanal *et al.*, 2017).

Plasmopara viticola, also known as oomycete grape downy mildew (DM), is a serious pathogen of grapevines that spreads through extremely efficient asexual reproduction cycles (Kiefer *et al.*, 2002). Originally from North America, DM was accidentally introduced to Europe at the end of the 19th century, where it caused extensive damage to the grape industry (Gessler *et al.*, 2011). The optimal environmental conditions for disease spread are high humidity and moderate temperatures. At early stages of infection, DM causes an increase in transpiration rate and a decrease in leaf temperature. Further DM development causes the appearance of chlorotic and necrotic tissue, increased water loss and the inability of plant tissue to regulate stomatal opening (Calderón *et al.*, 2014). IRT was used to detect spread of rose downy mildew infections one or two days before the appearance of visible symptoms (Caro, 2014) and cucumber downy mildew before visual symptoms as well (Wen *et al.*, 2019). The initial signs of infection in the thermal images were observed as early as 4 days after infection.

However, IRT is often subject to environmental factors such as ambient temperature, sunlight, rainfall, or wind speed (Mahlein, 2016). Changes in environmental conditions may affect leaf temperature, making it difficult to differentiate it from a change caused by infection or disease (Grant *et al.*, 2006). Alchanatis *et al.* (2010) found that for estimating and mapping water status variability of cotton, best results from thermal images were achieved at midday (12:00-14:00).

A previous study showed the feasibility of detecting grapevine downy mildew by SVM on a limited set of data (Cohen *et al.*, 2021). Looking at a small number of infection intervals, downy mildew was detected with an accuracy of 69.2% and F1 score of 74.9%

This study aimed to develop an algorithm for early detection of grapevine downy mildew using thermal imaging. The specific objectives were to: (i) extract features for classification based on temperature and image processing algorithms; (ii) develop classification models to classify between infected and healthy grapevine leaves; (iii) determine the best time along the day to acquire thermal images for downy mildew detection.

Materials and methods

Plant material and experimental design

Controlled experiments were conducted in six campaigns between the end of December 2019 and the end of October 2020 on 169 grapevine plants grown in experimental greenhouses in Evogene, Israel (31°47' 20.472" N, 35°12' 3.888" E). Each campaign lasted 12 days and included different imaging days (healthy leaves and 1, 2, 4, 5, 6, and 7 days after infection) in which around 30 plants were tested each day. Plants received two daily doses of water. Each of the six campaigns included the following stages:

1. The leaves that were selected to be infected were marked by color clips or aluminum foil.
2. On the first day of the campaign, images of the healthy leaves were acquired.
3. 2-6 leaves in each plant were infected with downy mildew and were inserted in a high humidity chamber which allowed optimal conditions for the disease to develop.

4. Images of each infected leaf were acquired on 1, 2, 4, 5, 6 and 7 days after infection.
5. After the last imaging day (day number 7), the leaves were placed in Petri dishes to evaluate the disease severity.

The diurnal response of leaf temperature was measured on October 27th. Images were acquired all day along at seven different times (rounds) between 7 a.m. and 4:30 p.m. (Table 1). Each round lasted about one and a half hours and included about 87 samples. In each plant, six leaves were acquired, of which three were infected with downy mildew six days after infection, and three were healthy. Each plant received two daily doses of water: one before the first round and one before the fourth round.

Table 1 Hours of rounds and number of samples in each round

Round number	Acquisition Time	Number of samples
1	7:15-8:25	88
2	9:00-9:45	88
3	10:40-11:30	88
4	12:25-13:15	87
5	14:15-15:05	86
6	15:20-16:00	86
7	16:00-16:30	52

Thermal and RGB image acquisition

To allow optimal photosynthesis of the plants, imaging was conducted outside of the greenhouse. The plants were taken out of the controlled greenhouse and placed outdoors for at least one hour before imaging to allow them to adjust to the environmental conditions that were different from greenhouse conditions. The images were acquired between 10 a.m. and 3 p.m. to ensure high solar radiation that allowed plants to conduct photosynthesis. Images were only acquired on sunny days without clouds. Meteorological conditions were continuously monitored and included measurements of air temperature (°C), relative humidity, solar radiation (W/m²), wind speed (m/sec), and wind direction. An image of each leaf was acquired by two cameras - a thermal camera (FLIR SC655, FLIR Systems, Melville, NY, USA) and an RGB camera (Canon EOS6D, Canon Inc., Taby, Sweden) that was used for documentation purposes. The infrared camera uses an uncooled microbolometer detector with resolution 640 × 480 pixels, sensitive in the spectral range of 7.5–13 μm and possesses an accuracy of ±2°C or ±2% of the reading. For each leaf, the thermal camera acquired a half-minute video and the RGB camera acquired two images. One image from each video was selected for classification. The image was manually selected by visually estimating the maximal leaf surface exposed to the camera.

Datasets

The classification dataset included 1403 records (599 healthy leaves and 804 infected leaves). The records included thermographic measurements from thermal imaging of the leaves, meteorological measurements collected simultaneously, calculated features from raw data, and manual evaluation of disease severity.

To determine the earliest day that a model can detect the disease, a subset was created from this dataset, which included records with actual disease severity of 5 or higher. This set

included 1097 records (599 healthy leaves and 498 infected leaves). Outliers were removed from the new set, which resulted in 1012 records (571 healthy leaves and 441 infected leaves).

The dataset for determining the best acquisition time contained 575 records (280 healthy leaves and 295 infected leaves).

Algorithm for leaves delineation

Leaves were segmented using edge detection with the 'Chan-Vese' active contour algorithm (Chan & Vese, 2001). This method ensures an unbiased contour enabling to either shrink or expand based on the image features. The software was implemented using MATLAB version R2019b (MathWorks Inc., Natick, MA, USA) with additional functions ([Shawn Lankton, 2007](#)). The inputs for the algorithm were a thermal image converted to grayscale and an initial mask; the outputs were an image with the leaf contour and a final mask. For each leaf, the position and size of the initial mask were set manually. The active contour algorithm was run on the mask for a maximum of 100 iterations with a smoothing term of 0.3 (Lambda).

Feature extraction

The following features were calculated from the leaf's mask: maximum temperature, minimum temperature, average temperature, median temperature, maximum temperature difference (MTD), standard deviation (STD), interquartile range (IQR), mean absolute deviation (MAD), Coefficient of variation (Cv), percentile 10, percentile 90 and Crop water stress index (CWSI).

All features expressing temperature were normalized by $T - T_{air}$.

Analysis

Outliers defined in this research as a data point that is 1.5 times the interquartile range above the upper quartile and below the lower quartile ($Q1 - 1.5 * IQR$ or $Q3 + 1.5 * IQR$) were removed. The correlation between the predictors was examined by Pearson's correlation coefficient. To avoid misleading information, the partial correlation coefficient between the predictors was also calculated. Using a correlation coefficient to determine whether there is a numerical relationship between two variables of interest will produce misleading results if there is another, confounding variable numerically related to them both. Therefore, the partial correlation coefficient controlling the confounding variable was used (Frank, 2000). The correlation between the predictors to the response variable was examined by point-biserial correlation.

To create the supervised learning model, the significance of the variables on the response variable was evaluated by stepwise logistic regression. The stepwise method combined "forward" and "backward" regression. The features selected by the stepwise logistic regression were used in all models. Pearson's correlation coefficient examined the correlation between the derived features.

To give equal importance to each feature and improve the model's accuracy, quality, and learning rate, the data were normalized using the Z-score (Dhaka et al., 2021). A zero centering of data was performed by subtracting the mean value from each attribute value, then dividing each dimension by its standard deviation. A Min-Max normalization was also used, but it did not significantly improve the results. Normalization and standardization were conducted with MATLAB.

In order to evaluate the statistical validity of the best model, statistical tests were conducted between the best model and the second-best model. An F-test was used to test the null hypothesis that the variances of both models are equal. A T-test was used to test the null hypothesis that both models have equal means.

A sensitivity analysis was conducted for different distributions of data to explain the different classification results between the days after infection.

Classification models

Five classification models were trained to classify infected and healthy leaves using MATLAB version R2019b (MathWorks Inc., Natick, MA, USA):

- Decision tree- one of the most widely used and practical methods for inference and classification. It has a fast prediction speed and is easy to interpret. This information gain method does not assume any statistical properties of the data itself (e.g., normal distribution) and as such, it is best suited to this case where the statistical distribution is unknown. When building a decision tree, overfitting may arise which is represented in the decision tree as a deep tree with many levels. To avoid over-fitting the maximum number of splits has been limited.
- Logistic regression- a statistical model that uses a logistic function to model a binary dependent variable and suitable in this case where there are two classes.
- Naïve bayes (NB)- a statistical classification technique based on Bayes Theorem. A simple supervised learning algorithm which provides fast and accurate classification. The classifier assumes that the effect of a particular feature in a class is independent of other features. However, the algorithm still appears to work well when the independence assumption is not valid.
- Support vector machine (SVM)- robust prediction models with very high accuracy of disease detection. An SVM training algorithm builds a model that assigns new examples to one category, making it a non-probabilistic binary linear classifier.
- Ensemble- The technique combines predictions from multiple machine-learning algorithms. In this work, the decision tree ensemble algorithm using the Boosting method was used. Boosting refers to a group of algorithms that trains weak learners sequentially, each trying to correct its predecessor.

K-fold cross-validation (K=5) was used for each model.

Performance measures

The classification performance was quantified using the accuracy, precision, recall, F1 score and AUC (Area under the ROC Curve).

The class set contained two labels: positive (infected) and negative (healthy). Given a classifier and an instance, there were four possible outcomes:

- True positive (TP): the leaf was infected and it was classified as infected.
- False-negative (FN): the leaf was infected, but it was classified as healthy.
- True negative (TN): the leaf was healthy and it was classified as healthy.
- False-positive (FP): the leaf was healthy, but it was classified as infected.

Accuracy is defined as the probability of correctly classifying a test instance:

$$Accuracy = \frac{TP + TN}{Total\ number\ of\ instances} \quad (1)$$

Precision is called positive predictive value and computed as:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Recall is also referred to true positive rate (TPR) and computed as:

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

The F1 score is the harmonic mean of the precision and recall and computed as:

$$F1\ score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

A ROC curve (receiver operating characteristic curve) is a graph showing the performance of a classification model at all classification thresholds displaying two parameters:

- Recall (also TPR)
- False Positive Rate (FPR)

False Positive Rate is defined as follows:

$$FPR = \frac{FP}{FP + TN} \quad (5)$$

A ROC curve plots TPR vs. FPR at different classification thresholds. Lowering the classification threshold classifies more items as positive, thus increasing both False Positives and True Positives.

AUC stands for "Area under the ROC Curve". The AUC measures the entire two-dimensional area underneath the entire ROC curve.

Results and Discussion

Edge detection

First, the thermal image was converted to grayscale. Second, the range of pixel values was changed by cutting the histogram edges to clear the color contrast. This change was made for visualization purposes only. Then, the active contour algorithm was activated on the image. Fig. 1 depicts the active contour process. Image and initial mask were the algorithm inputs (Fig. 1A(a), 1A(b)). The algorithm performed iterations to find the contours of the leaf (Fig. 1A(c)). After a maximum of 100 iterations a final mask was obtained (Fig. 1A(d)). An image of the leaf with contour was returned (Fig. 1B). The returned final mask was a matrix where pixels inside the mask were set to leaf temperature and pixels outside the mask were set to zero.

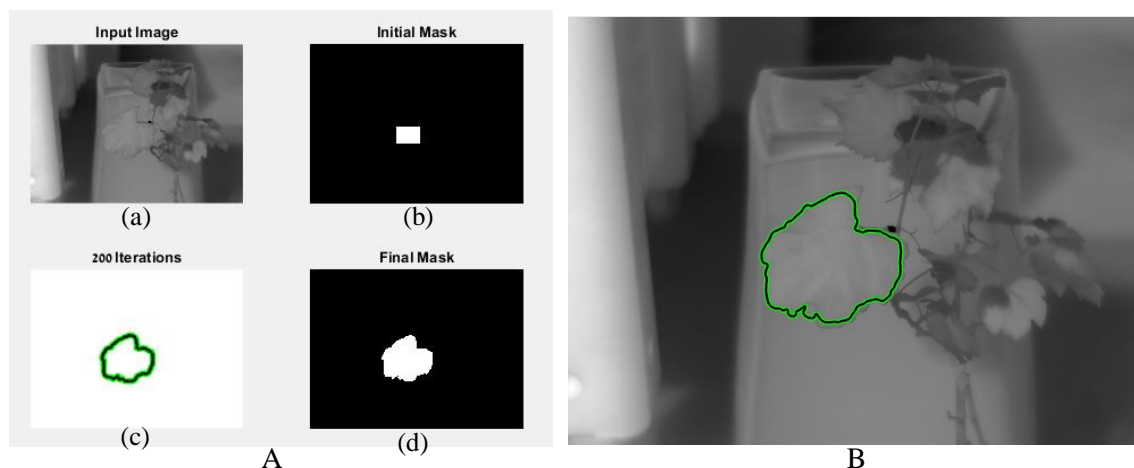


Fig. 1 The output of the active contour algorithm. Includes image, initial and final mask (A) and image with final mask (B)

Classification of healthy and infected leaves

Fig. 2 depicts examples of thermal and RGB images acquired at different days of infection. Compared with the digital images, the changes in the color of the typical thermal images were more noticeable compared to the visual observations.

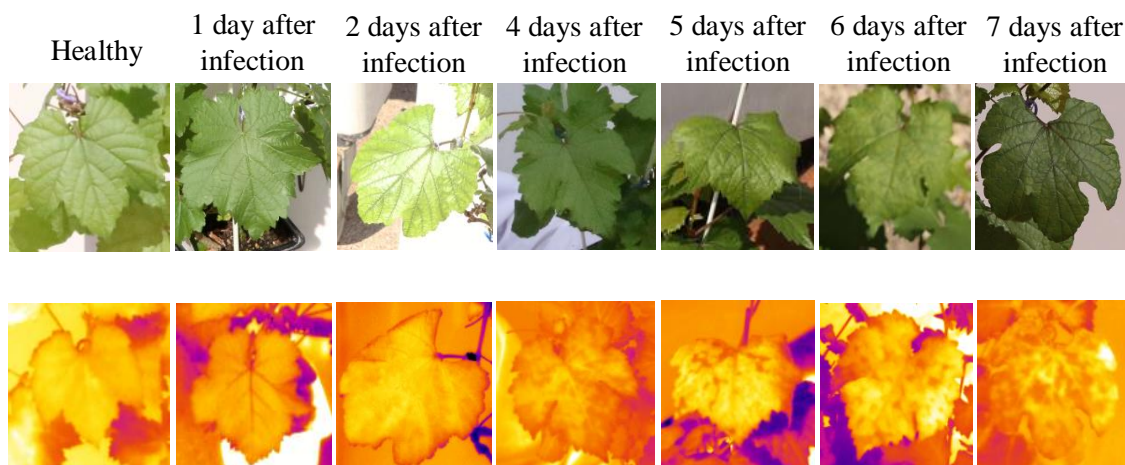


Fig. 2 Examples of RGB images (top row) and thermal images (bottom row) of leaves from different infected days

Feature selection

The feature selection conducted by stepwise logistic regression (Table 2) resulted in seven significant variables: MTD, STD, Cv, T_{avg} , median temperature, percentile 90, CWSI. This result was unexpected because Pearson's partial correlation matrix showed that T_{avg} and median temperature were in a high positive linear correlation ($r=0.97$), which led to the expectation that the model would not include both. Another unexpected result was the estimate of T_{avg} ($\beta=-9.53$) which had the opposite sign of the median temperature ($\beta=3.696$). A possible reason for these results is that the selected method for stepwise logistic regression ("both") begins with the full model. When the two variables (T_{avg} and median temperature) are together in the model and have high multicollinearity and opposite sign, they cancel each other out and stay in the model instead of removing both. Therefore, both variables were

removed from the model manually. This led to a higher AIC, resulting with a better model. Finally, to train the classification models the five remained significant variables were selected: MTD, STD, Cv, percentile 90, CWSI.

Table 2 Stepwise regression- estimated coefficients, standard errors and p-value

Variable	Estimated coefficients	Standard errors	P-value
MTD	0.6543	0.1782	0.00024
STD	-2.4373	1.1323	0.03136
Cv	79.2226	24.7641	0.00138
percentile 90	0.2709	0.0388	3.06e-12
CWSI	1.6405	0.4575	0.00034

Classification analysis

The parameters for each model were optimized using MATLAB version R2019b. The optimal parameters obtained for each model were:

- Decision Tree: 17 maximum splits, Maximum deviance reduction split criterion
- Logistic Regression: Threshold 0.5
- Naive Bayes: Gaussian kernel
- SVM: Cubic Kernel, box constraint level of 1
- Ensemble: LogitBoost method, 14 maximum splits, a learning rate of 0.348

The accuracy, precision, recall, F1 score and AUC of all constructed models (Table 3) reveal that the best results were achieved by the SVM model with major differences between the performance measures of the other models.

The results of the F-test between the best model (SVM) and the second-best model (NB) using the F1 scores of each of the five folds (cross-validation) revealed that the variances of both models were equal ($F (=3.27) < F$ Critical one-tail ($=6.39$)) implying that the null hypothesis was not rejected).

The two-tail T-test assuming equal variances revealed that the null hypothesis was rejected (t Stat ($=-5.56$) $< t$ Critical two-tail ($=-2.31$)). The observed difference between the sample means (0.77 - 0.67) suggests a significant difference between the models. SVM was chosen to be the best model.

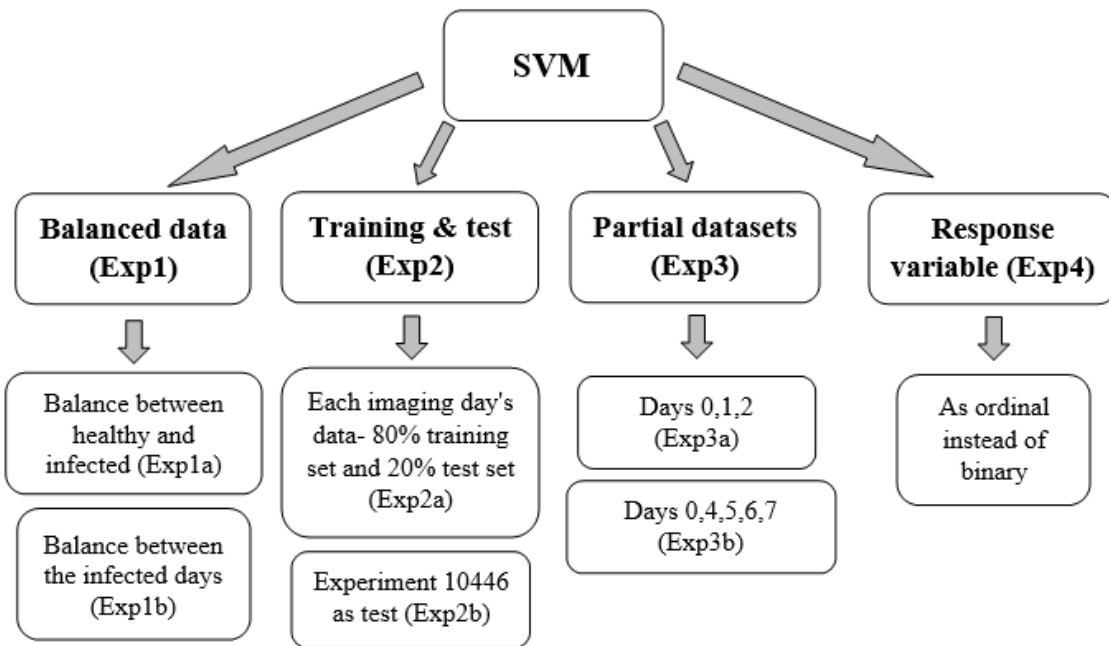
To determine the earliest day for downy mildew detection, the results in the following table (Table 4) obtained by the SVM were detailed according to each of the 'days after infection'.

Table 3 All results from all models based on all data

Model	Decision Tree	Logistic Regression	NB	SVM	Ensemble
Measure					
F1 score	60.5%	64.9%	66.8%	77.5%	63.8%
Precision	70.5%	70.8%	69.7%	83.1%	70.9%
Recall	53.1%	59.9%	64.2%	71.6%	58%
AUC	0.728	0.762	0.783	0.874	0.776
Accuracy	69.9%	71.7%	72.2%	81.6%	71.3%

Table 4 Results by day after infection

Days after infection	Number of samples	Number of misses	Accuracy
0	571	65	88.6%
1	19	2	89.5%
2	61	3	95.1%
4	180	55	69.4%
5	39	4	89.7%
6	44	17	61.4%
7	98	40	59.2%
Model			81.6%

**Fig. 3** Diagram representing the logical flow of the different analyses

The results revealed that later days after infection produced lower results than earlier ones, contrary to expectations (since the disease develops over time and should be easier to detect as time passes). To determine the cause of this, a series of analyses as described in Fig. 3, were conducted based on different explanations that may have led to these results.

1. To avoid bias, the dataset was balanced.
 - a. By removing records from the healthy leaves, the samples number between the diseased and healthy leaves was balanced. The accuracy achieved with these balance results was 79.1%, the F1 score was 77.9%, and the AUC was 0.86. The model's accuracy of the later days after infection improved (Table 5).

Table 5 Results after balancing healthy and infected samples (Exp1a)

Days after infection	Number of samples	Number of misses	Accuracy
0	441	68	84.6%
1	19	1	94.7%
2	61	5	91.8%
4	180	55	69.4%
5	39	5	87.2%
6	44	14	68.2%
7	98	36	63.3%
Model			79.1%

- b. Besides taking similar numbers of samples from healthy and infected leaves, the number of samples taken each day after infection was also balanced. The accuracy achieved was 73.8%, the F1 score was 71%, and the AUC was 0.756. This did not yield improvement in the model's results, but the accuracy of the later days after infection improved greatly.
2. To examine the effects of different climatic conditions on the results, the data was divided into different training and test sets.
 - a. As the climatic conditions differed between imaging days, it was difficult to classify the data. Each imaging day's data was split in two: 80% from the data for the training set and 20% from the data for the test set. The accuracy of the training set was 82.5%, with an F1 score of 78.3% and an AUC of 0.886. The test set accuracy was 76.5%, with an F1 score of 70.8% and an AUC of 0.827. Some days' results improved, while others did not. The model's performance did not improve.
 - b. The test set included one specific experiment (No. 10446), and the training set included the rest. Experiment 10446 included the days 0,4,5,6,7 after infection. Days 5,6 were not included in any other experiments, so they appear only in the test set. The accuracy of the training set was 86.4%, the F1 score was 82%, and the AUC was 0.892. The test set accuracy was 57.8%, with an F1 score of 48.9% and an AUC of 0.593. The training and test sets were very different in their accuracy. Even on the days that appear in the training set (0,4,7), the results are poor. According to this analysis, it is not possible to classify healthy and downy mildew infected from images acquired on an imaging day that is not included in the training set.
3. Considering that the accuracy on days 1 and 2 after infection was greater than that of the other days, it was tested whether these days affect the prediction.
 - a. The following hypothesis was tested: whether with a dataset of leaves from healthy and 1 and 2 days after infection, the accuracy of the prediction for these days still remains high (Table 6). The data for this analysis was balanced. The accuracy achieved was 91.9%, the F1 score was 92.1%, and the AUC was 0.961. The results showed that even without the later days, the early days' predictions held true very well.

Table 6 Results of a dataset with days 0,1,2 after infection (Exp3a)

Days after infection	Number of samples	Number of misses	Accuracy
0	99	10	89.9%
1	21	1	95.2%
2	78	5	93.6%
Model			91.9%

- b. The same was done by using a dataset of days 0,4,5,6,7 after infection. Here also the data was balanced (Table 7). The accuracy achieved was 79.1%, the F1 score was 78.4%, and the AUC was 0.856. The results showed that when days 1 and 2 after infection were removed, the classification results of the later days improved.

Table 7 Results of a dataset with days 0,4,5,6,7 after infection (Exp3b)

Days after infection	Number of samples	Number of misses	Accuracy
0	399	72	81.9%
4	197	54	72.6%
5	45	5	88.9%
6	45	9	80%
7	112	27	75.9%
Model			79.1%

4. Since the results of the other approaches varied each day after infection, an analysis was performed to examine if the response variable should be ordinal instead of binary. The assumption was that disease development increased every day, suggesting some kind of order as imaging days progressed. Therefore, an ordinal regression was performed on the imbalanced data. A new feature selection was conducted to accommodate for the new response variable and resulted in the following features: MTD, IQR, MAD, median, perc10, perc90, and CV. The accuracy of the ordinal classification was 65.4% with many observations classified as healthy even though they were infected (and were not classified as infected on another day, Fig. 4). Since it does not matter which day after infection was classified, but whether the leaf was healthy or infected, the results were converted to binary so all days that were not classified as 0 (healthy) were deemed infected. This classification of the converted response variable resulted in an accuracy of 74.9%. The binary response variable performed better than both the ordinal response variable and the converted response variable and hence was selected.

0	510	6	6	28	2	5	14
1	10	2	3	3			1
2	12	6	33	5			5
4	86	1	2	75	6	4	6
5	13			6	6	9	5
6	23			6	3	7	5
7	49		2	11	3	4	29
	0	1	2	4	5	6	7

Figure 4 Confusion matrix of ordinal regression

All approaches were also evaluated after a new feature selection process was performed (except for Exp4, which anyway included a new feature selection). The purpose of this was to examine whether the results and features selected were different. In each approach, the new selected features were very similar to the first selection, with performance differing by only $\pm 3\%$. These results indicated that the first feature selection was indeed suitable for the different datasets.

Table 8 summarizes the results of the different approaches used to explain the differences in disease detection between days after infection. Although the best results were with an SVM model using a dataset with healthy and infected leaves from only 1 and 2 days after infection, this dataset itself is not selected since it included only the two first days after infection without the remaining days. These good results were probably due to the fact that the plants were disturbed by the controlled infection which was what was probably actually detected. The best relevant results were obtained from a SVM model using a balanced dataset of healthy leaves and infected leaves.

A possible explanation for the differences in disease detection between days after infection might be that infrared thermography measurements are influenced by many other factors that affect leaf temperature, including ambient temperature, humidity, sunlight, and wind (Granum et al., 2015). The images were acquired at different dates, times of day, and under different environmental conditions, which may have affected the characteristics of the images or even the plants. Although each variable related to temperature has been normalized ($T-T_{air}$), this might not be sufficient to accommodate environmental changes.

The classification results of days one and two after infection, which were better than those of the other days, was probably not because the downy mildew was detected early; It is more likely that the infection itself, which was manually induced, may have affected the leaves locally and caused a strong physiological reaction to the leaves. As the virus penetrated the plant, the local effect diminished.

Despite many different approaches and analyses conducted, we could not provide an explanation for the inconsistent results between the days (days 4-7 after infection).

Since each of the models (the model with data of days 0,1,2 and the model with data of days 0,4,5,6,7) provided good results, future research should consider dividing the data into 3 groups: healthy, infected at first days, and infected at late days.

Table 8 A summary of all approaches and their results

Approach	F1 score	AUC	Accuracy
SVM- all data	77.5%	0.874	81.6%
Balance between healthy and infected (Exp1a)	77.9%	0.86	79.1%
Balance between the infected days (Exp1b)	71%	0.756	73.8%
Each imaging day's data- 80% training set and 20% test set (Exp2a)	70.8%	0.827	76.5%
Experiment 10446 as test (Exp2b)	48.9%	0.593	57.8%
Days 0,1,2 (Exp3a)	92.1%	0.961	91.9%
Days 0,4,5,6,7 (Exp3b)	78.4%	0.856	79.1%
As ordinal instead of binary (Exp4)			74.9%

Best acquisition time

Examples of thermal images from infected and healthy leaves acquired at different times of the day are depicted in Fig. 4. The displayed images have the same scale (23.6°C - 34.6°C).

Table 9 presents accuracy, F1 score and AUC obtained by SVM model for all rounds. Best results were obtained when data were collected 10:40-11:30 (round 3) with accuracy differences of about 6% and 23% from the next best and worst rounds respectively. These results are different from results of Alchanatis et al. (2010), who found that midday (12:00-14:00) was the best time to map and estimate water status variability using thermal imaging. Although these are two different problems, they are similar in expectance to see temperature differences within the leaves.

The results of the F-test between the best round (3) and the second-best round (1) using the accuracy of each of the five folds from cross-validation (since the data was balanced) revealed that the null hypothesis was not rejected ($F (=4.08) < F \text{ Critical one-tail } (=6.39)$), implying that the variances of both rounds were equal.

The two-tail T-test assuming equal variances revealed that the null hypothesis was not rejected ($-t \text{ Critical two-tail } (= -2.31) < t \text{ Stat } (= -1.12) < t \text{ Critical two-tail } (= 2.31)$). The observed difference between the sample means (0.807 - 0.751) is not convincing enough to say that the results differ significantly.

Although the 3rd round was not statistically better, records from these hours were analyzed to generalize the findings since best results were obtained at this acquisition time. A special new dataset of 239 records (95 healthy and 144 infected) was created using only observations that were acquired 10:40 and 11:30 a.m. from the classification model (the complete dataset). This dataset was used to create a new model; the results are detailed in Table 10. To improve the results, a new feature selection was performed. The selected features were: MTD, Tavg, Tmin, perc10, perc90, and Cv. The new features improved the results and the model produced good results (Table 10).

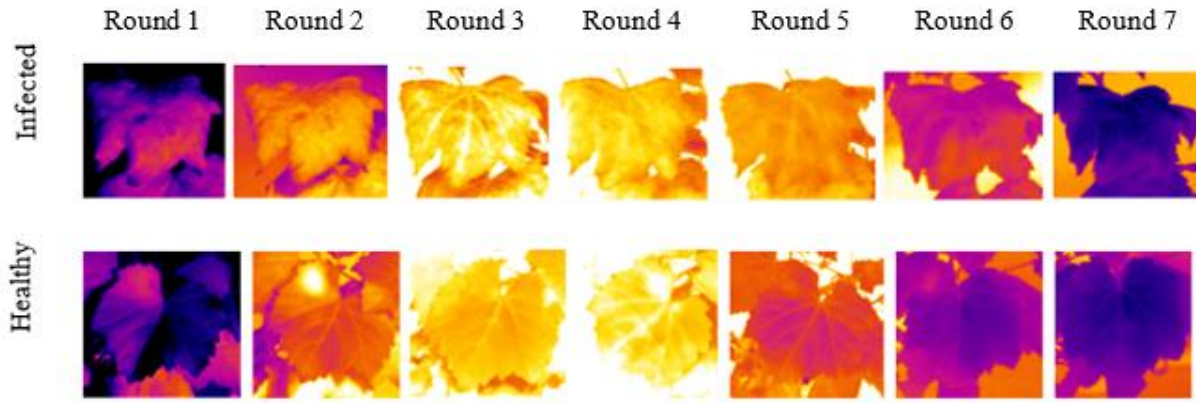


Fig. 4 Thermal images from infected and healthy leaves acquired at different times of the day

Table 9 Performance for each round of the diurnal measurements

Round No./Time	Measure	Accuracy	F1 score	AUC
(1) 7:15-8:25		75%	75.6%	0.774
(2) 9:00-9:45		72.7%	72.7%	0.794
(3) 10:40-11:30		80.7%	80.5%	0.895
(4) 12:25-13:15		59.8%	61.5%	0.676
(5) 14:15-15:05		65.1%	67.4%	0.691
(6) 15:20-16:00		58.1%	61.7%	0.644
(7) 16:00-16:30		57.7%	59.3%	0.557

Table 10 Results of a dataset from 10:40 to 11:30 and after a new feature selection

Model	Number of samples	F1 score	AUC	Accuracy
Hours 10:40-11:30	239	72.7%	0.764	67.4%
New features	239	80.8%	0.826	76.6%

As compared with the model that used all the data, the model from the best hours included fewer observations, but the results were quite similar. This model would likely perform even better if it were trained on a larger dataset. Therefore, it is worth considering acquiring images at similar times along the day, even if they are acquired on different days. Further research is needed on this topic.

Though, even when the acquisition occurred during the best hours, results are strongly affected by the environmental conditions, which may change day by day, and impact the results.

Conclusions

The best model for classifying between healthy and infected leaves was a SVM model built on a balanced dataset with the following features: MTD, STD, percentile 90, Cv and CWSI. The model resulted in 10% higher performance than all other models tested (81.6% accuracy,

77.5% F1 score and 0.874 AUC). The inconsistent results between the days (days 4-7 after infection) could not be explained.

The best time of day for acquiring images for downy mildew detection was between 10:40 and 11:30 resulting in 80.7% accuracy, 80.5% F1 score, and 0.895 AUC.

The results indicate that thermograms can detect downy mildew infection, even before any visible symptoms appear. Using images from the best hours can improve performance, even if the images are not from the same days.

However, even when the image acquisition is conducted at the best time, variations in illumination cannot be avoided, which results in reduced performance. There is a trade-off between using a large and wide database (acquired along many dates) and detecting the disease (the fewer dates, the easier).

Acknowledgements

We wish to thank Mr. S. Ozer for his support during the image acquisition campaigns. We are also grateful to Z. Halamish, N. Guedj, A. Himmel and A. Movchan for their excellent support on growing the plants and on conducting the experiments.

Declarations

Funding: Financial support was provided by the Israel Innovation Authority, project No. 67268. Partial financial support was received from Ben-Gurion University of the Negev through the Rabbi W. Gunther Plaut Chair in Manufacturing Engineering.

Conflicts of interest/Competing interests: The authors declare that they have no conflict of interest.

Availability of data and material:

Data sets (Excel): <https://github.com/BarCohenBGU/database.git>

'All data new'- the classification dataset included 1403 records

'normalized_by_severity'- included 1097 records with actual disease severity of 5 or higher

'features_without_outliers'- included 1012 records after removing outlier

Daily dataset (Excel): <https://github.com/BarCohenBGU/daily-dataset.git>

Code availability:

Active contour algorithm (MATLAB): <https://github.com/BarCohenBGU/active-contour-algorithm.git>

Histograms and boxplots, Outliers, Correlation (Python): <https://github.com/BarCohenBGU/statistics.git>

Stepwise logistic regression (RStudio): <https://github.com/BarCohenBGU/feature-selection.git>

Classification models (MATLAB): <https://github.com/BarCohenBGU/classification-models.git>

References

Alchanatis, V., Cohen, Y., Cohen, S., Moller, M., Sprinstin, M., Meron, M., Tsipris, J., Saranga, Y., & Sela, E. (2010). Evaluation of different approaches for estimating and mapping crop water status in cotton with thermal imaging. *Precision Agriculture*, 11(1), 27–41. <https://doi.org/10.1007/s11119-009-9111-7>

- Calderón, R., Montes-Borrego, M., Landa, B. B., Navas-Cortés, J. A., & Zarco-Tejada, P. J. (2014). Detection of downy mildew of opium poppy using high-resolution multi-spectral and thermal imagery acquired with an unmanned aerial vehicle. *Precision Agriculture*, 15(6), 639–661. <https://doi.org/10.1007/s11119-014-9360-y>
- Caro, S. G. (2014). Infection and spread of *Peronospora sparsa* on *Rosa* sp.(Berk.) - a microscopic and a thermographic approach. *Doctoral Dissertation, University of Bonn, Germany*.
- Chan, T. F., & Vese, L. A. (2001). Active contours without edges. *IEEE Transactions on Image Processing*, 10(2), 266–277. <https://doi.org/10.1109/83.902291>
- Cohen, B., Edan, Y., Levi, A., & Alchanatis, V. (2021). Early detection of grapevine downy mildew using thermal imaging. In *Precision agriculture'21* (pp. 283-290). Wageningen Academic Publishers. <https://doi.org/10.3920/978-90-8686-916-9>
- Cui, S., Ling, P., Zhu, H., & Keener, H. M. (2018). Plant pest detection using an artificial nose system: A review. *Sensors (Switzerland)*, 18(2), 1–18. <https://doi.org/10.3390/s18020378>
- Dhaka, V. S., Meena, S. V., Rani, G., Sinwar, D., Kavita, Ijaz, M. F., & Woźniak, M. (2021). A survey of deep convolutional neural networks applied for prediction of plant leaf diseases. *Sensors*, 21(14), 4749. <https://doi.org/10.3390/s21144749>
- Frank, K. A. (2000). impact of a confounding variable on a regression coefficient. *Sociological Methods & Research*, 29(2), 147–194.
- Gessler, C., Pertot, I., & Perazzolli, M. (2011). *Plasmopara viticola*: A review of knowledge on downy mildew of grapevine and effective disease management. *Phytopathologia Mediterranea*, 50(1), 3–44. https://doi.org/10.14601/Phytopathol_Mediterr-9360
- Grant, O. M., Chaves, M. M., & Jones, H. G. (2006). Optimizing thermal imaging as a technique for detecting stomatal closure induced by drought stress under greenhouse conditions. *Physiologia Plantarum*, 127(3), 507–518. <https://doi.org/10.1111/j.1399-3054.2006.00686.x>
- Granum, E., Pérez-Bueno, M. L., Calderón, C. E., Ramos, C., de Vicente, A., Cazorla, F. M., & Barón, M. (2015). Metabolic responses of avocado plants to stress induced by *Rosellinia necatrix* analysed by fluorescence and thermal imaging. *European Journal of Plant Pathology*, 142(3), 625–632. <https://doi.org/10.1007/s10658-015-0640-9>
- Khanal, S., Fulton, J., & Shearer, S. (2017). An overview of current and potential applications of thermal remote sensing in precision agriculture. *Computers and Electronics in Agriculture*, 139, 22–32. <https://doi.org/10.1016/j.compag.2017.05.001>
- Kiefer, B., Riemann, M., Büche, C., Kassemeyer, H. H., & Nick, P. (2002). The host guides morphogenesis and stomatal targeting in the grapevine pathogen *Plasmopara viticola*. *Planta*, 215(3), 387–393. <https://doi.org/10.1007/s00425-002-0760-2>
- Mahlein, A.-K. (2016). Plant disease detection by imaging sensors- parallels and specific demands for precision agriculture and plant phenotyping. *Plant Disease*, 100(2), 1–11. <https://doi.org/10.1007/s13398-014-0173-7.2>
- Sankaran, S., Mishra, A., Ehsani, R., & Davis, C. (2010). A review of advanced techniques for detecting plant diseases. *Computers and Electronics in Agriculture*, 72(1), 1–13. <https://doi.org/10.1016/j.compag.2010.02.007>
- Savary, S., Ficke, A., Aubertot, J. N., & Hollier, C. (2012). Crop losses due to diseases and their implications for global food production losses and food security. *Food Security*, 4(4), 519–537. <https://doi.org/10.1007/s12571-012-0200-5>
- Wen, D. M., Chen, M. X., Zhao, L., Ji, T., Li, M., & Yang, X. T. (2019). Use of thermal imaging and Fourier transform infrared spectroscopy for the pre-symptomatic detection of cucumber downy mildew. *European Journal of Plant Pathology*, 155(2), 405–416. <https://doi.org/10.1007/s10658-019-01775-2>

Appendix B – ECPA conference paper

Early detection of grapevine downy mildew using thermal imaging

B. Cohen^{1,2}, Y. Edan², A. Levi¹, V. Alchanatis¹

¹Agricultural Research Organization, Agricultural Engineering, Rishon Le Zion, P.O. Box 15159, 7528809, Israel

²Ben-Gurion University of the Negev, Industrial Engineering and Management, Beer-Sheva, P.O. Box 653, 84105, Israel
victor@volcani.agri.gov.il

Abstract

The objective of this research was to detect downy mildew in grapevine leaves at early stages of development using thermal imaging technology based on the assumption that plant disease causes significant modifications in leaf temperature. Infected and healthy leaves of grapevine grown in a controlled greenhouse experiment were classified using thermal images that were acquired 1, 2, 4 and 7 days from infection. Leaves were detected using the active contour algorithm for edge detection. The resulting leaf mask was used to calculate 14 features. Stepwise regression analysis revealed four significant features that were consequently used in prediction models. Five models were developed to classify between infected and healthy leaves. The best results were obtained by the support vector machine model using cross-validation on all data with a classification accuracy of 69.2% and an F1 score of 74.9%. Identifying downy mildew 4 and 7 days after infection resulted in an accuracy of 66.4% and 83.1% respectively.

Keywords: Thermal imaging, Downy mildew, Grapevine, Disease detection

Introduction

Plant diseases are a significant factor causing production and economic losses in the agriculture industry (Savary *et al.*, 2012). Direct yield losses of 20-40% of global agricultural productivity are caused by pathogens (Savary *et al.*, 2012). Therefore, it is critical for sustainable agriculture to monitor plant health to prevent disease spread with the least damage to crop production. However, the main challenge is the difficulty in determining the physical, chemical and biological changes in plants, before there are symptoms of an infection (Cui *et al.*, 2018).

Disease detection techniques can be classified into invasive and non-invasive methods. Invasive techniques involve destructive leaf sampling followed by chemical treatments after direct identification of the pathogen (Sankaran *et al.*, 2010). Non-invasive techniques identify plant diseases by detecting the impact of the pathogen on the physiological plant response. Currently, the most promising non-invasive disease detection techniques are sensors that measure temperature, reflectance or fluorescence (Mahlein, 2016).

A rapid response variable that can indicate crop stresses before visual symptoms appear is leaf temperature (Khanal *et al.*, 2017). A plant that is in stressed conditions reacts with physiological protection mechanisms leading to changes in leaf surface temperature (A.-K. Mahlein, 2016). Infrared thermography (IRT) enables to evaluate plant temperature that is related to plant water status and changes in transpiration as a result of infections by plant pathogens. The temperature difference within single leaves, plants and crop stands is a strong indicator of the appearance of plant disease (Mahlein, 2016). Thermal sensors have shown to

be more effective in detecting disease-induced early modifications compared to optical, multi and hyperspectral sensors (Khanal *et al.*, 2017).

The oomycete *Plasmopara viticola*, called also oomycete grape downy mildew (DM), is a serious pathogen of grapevine and spreads by extremely efficient cycles of asexual propagation (Kiefer *et al.*, 2002). DM is native to North America and was accidentally introduced into Europe at the end of the 19th century, where it caused widespread damage to the grape industry (Gessler *et al.*, 2011). The optimal environmental conditions for disease spread are high humidity and moderate temperatures. Infection by DM causes an increase in transpiration rate and a decrease in leaf temperature at early stages of infection. Further DM development causes the appearance of chlorotic and necrotic tissue, increased water loss and the inability of plant tissue to regulate stomatal opening (Calderón *et al.*, 2014). Caro (2014) showed that IRT confirmed the spread of rose downy mildew infections. This technique detected the disease on roses one or two days before the appearance of visible symptoms. Wen *et al.* (2019) found that thermal infrared imaging could identify cucumber downy mildew before visual symptoms. The initial signs of infection in the thermal images were observed as early as 4 days after infection.

This study aimed to develop an algorithm for early detection of grapevine downy mildew using thermal imaging. The specific objectives were to develop: (i) an image processing algorithm to derive the leaf outline and determine temperature parameters; (ii) a model to classify between infected and healthy grapevine leaves.

Materials and methods

Plant material

Controlled experiments including treatments and imaging were undertaken in five campaigns and included 154 grapevine plants. The experiments took place from the end of December 2019 to early April 2020 in experimental greenhouses of Evogene, Israel (31° 47' 20.472" N, 35° 12' 3.888" E). Each campaign lasted 12 days and included 5 imaging days (healthy leaves and 1, 2, 4 and 7 days after infection) in which around 30 plants were tested each day. Each plant received two daily doses of water. In each of the five campaigns, the following stages were carried out:

6. The leaves that were selected to be infected were tagged by color clips.
7. On the first day of the campaign, images of the healthy leaves were acquired.
8. 2-6 leaves in each plant were infected with downy mildew and were inserted in a high humidity chamber which allowed optimal conditions for the disease to develop.
9. Imaging of each infected leaf on 1, 2, 4 and 7 days after infection.

Thermal and RGB image acquisition

To allow the plants to photosynthesize optimally, imaging was conducted outside of the greenhouse. The plants were placed outside at least one hour before imaging to adjust to the environmental conditions that were different from greenhouse conditions. The images were acquired between 10 a.m. and 3 p.m. to ensure high solar radiation that allowed plants to photosynthesize. Images were acquired only on sunny days with no clouds. Meteorological conditions were continuously acquired and included measurements of air temperature (°C), relative humidity, solar radiation (W/m²), wind speed (m/sec) and wind direction. An image of each leaf was acquired by two cameras - a thermal camera (FLIR SC655, FLIR Systems, Melville, NY, USA) and an RGB camera (Canon EOS6D, Canon Inc., Taby, Sweden) that was used for documentation purposes. The infrared camera uses an uncooled microbolometer detector with resolution 640 × 480 pixels, sensitive in the spectral range of 7.5–13 μm and possesses an accuracy of ±2°C or ±2% of the reading. For each leaf, the thermal camera acquired a half-minute video and the RGB camera acquired two images. One image from each

video was selected for classification. The image was manually selected by visually estimating the maximal leaf surface.

Database

The database included 1065 records (425 healthy leaves and 640 infected leaves). The records included thermographic measurements from the thermal imaging of the leaves and environmental measurements from the meteorological station that were acquired simultaneously.

To determine the earliest day a model can detect the disease, the data were divided into sets where each set contained different days of infected leaves:

- Set 1- 1065 records; 640 leaves from 1,2,4,7 days after infection and 425 healthy leaves.
- Set 2- 994 records; 569 leaves from 2,4,7 days after infection and 425 healthy leaves.
- Set 3- 826 records; 401 leaves from 4,7 days after infection and 425 healthy leaves.

Algorithm for leaves delineation

Leaves were detected using edge detection with the 'Chan-Vese' active contour algorithm (Chan & Vese, 2001). This method ensures an unbiased contour enabling to either shrink or expand based on the image features. The software was implemented using MATLAB version R2019b (MathWorks Inc., Natick, MA, USA) with additional functions ([Shawn Lankton, 2007](#)). The inputs for the algorithm were a thermal image converted to grayscale and an initial mask; the outputs were an image with the leaf contour and a final mask. For each leaf, the position and size of the initial mask were set manually. The active contour algorithm was run on the mask for a maximum of 200 iterations with a Lambda parameter (the smoothing term) of 0.3.

Feature extraction and selection

The following features were calculated from the leaf's mask: maximum temperature, minimum temperature, average temperature, maximum temperature difference (MTD), standard deviation (STD), median, percentile 2, 10, 25, 75, 90, 98, interquartile range (IQR) and mean absolute deviation (MAD). All features expressing temperature were normalized by $T - T_{air}$ and replaced the non-normalized features.

To create a supervised learning model, the significance of the variables on the response variable was evaluated by stepwise regression (stepwisefit, Statistics and Machine Learning Toolbox, MATLAB) . A variable with a low significance ($p > 0.1$) was removed until the most significant variables were obtained. Since there is no common method for feature selection, the features selected by stepwise regression were used in all models. Pearson's correlation coefficient examined the correlation between the derived features.

Classification models

Five classification models were trained to classify infected and healthy leaves using MATLAB version R2019b (MathWorks Inc., Natick, MA, USA):

- Decision tree- one of the most widely used and practical methods for inference and classification. It has a fast prediction speed and is easy to interpret. This information gain method does not assume any statistical properties of the data itself (e.g., normal distribution) and as such, it is best suited to this case where the statistical distribution is unknown. When building a decision tree, overfitting may arise which is represented in the decision tree as a deep tree with many levels. To avoid over-fitting the maximum number of splits has been limited.
- Logistic regression- a statistical model that uses a logistic function to model a binary dependent variable and suitable in this case where there are two classes.

- Naïve bayes (NB)- a statistical classification technique based on Bayes Theorem. It is one of the simplest supervised learning algorithms. NB classifier is a fast, accurate and reliable algorithm. The classifier assumes that the effect of a particular feature in a class is independent of other features. However, the algorithm still appears to work well when the independence assumption is not valid.
- Support vector machine (SVM)- robust prediction models with very high accuracy of disease detection. An SVM training algorithm builds a model that assigns new examples to one category, making it a non-probabilistic binary linear classifier.
- Ensemble- The technique combines predictions from multiple machine-learning algorithms. In this work, the decision tree ensemble algorithm using the Boosting method was used. Boosting refers to a group of algorithms that trains weak learners sequentially, each trying to correct its predecessor.

The classification was performed by K-fold cross-validation (K=10) to avoid model overfitting.

Performance measures

The performance of classification was quantified using accuracy, precision, recall and F1 score. The class set contained two labels: positive (infected) and negative (healthy). Given a classifier and an instance, there were four possible outcomes:

- True positive (TP): the leaf was infected and it was classified as infected.
- False-negative (FN): the leaf was infected, but it was classified as healthy.
- True negative (TN): the leaf was healthy and it was classified as healthy.
- False-positive (FP): the leaf was healthy, but it was classified as infected.

Accuracy is defined as the probability of correctly classifying a test instance:

$$Accuracy = \frac{TP + TN}{Total\ number\ of\ instances} \quad (1)$$

Precision is called positive predictive value and computed as:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Recall is also referred to true positive rate and computed as:

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

The F1 score is the harmonic mean of the precision and recall and computed as:

$$F1\ score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

Results and Discussion

Figure 1 depicts examples of thermal and RGB images acquired at different days of infection. More hot-spots appear in the thermal image of the leaf as more days have passed since the infection date.

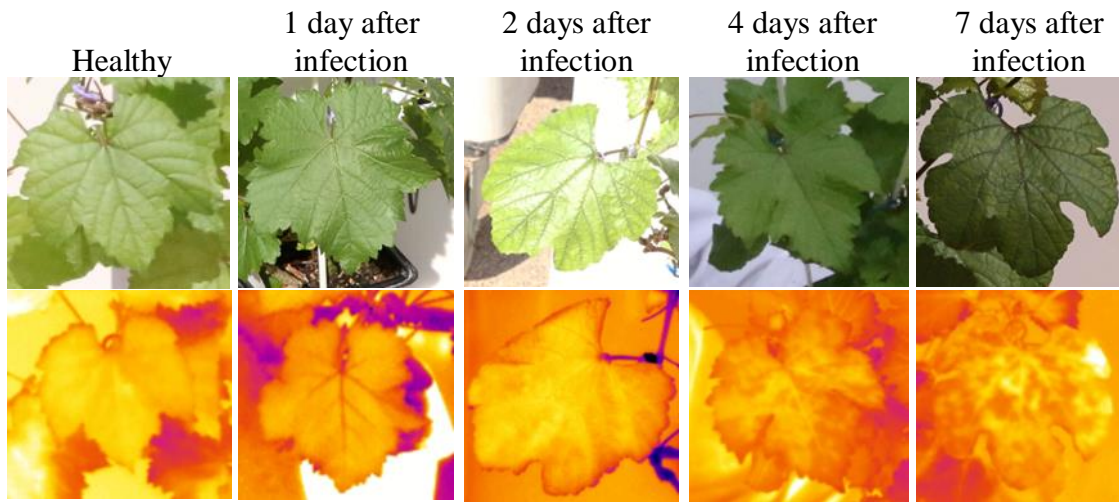


Figure 1. Examples of RGB images (top row) and thermal images (bottom row) of leaves from different infected days

Figure 2 depicts the active contour process. Image and initial mask were the algorithm inputs (Fig. 2A(a), 2A(b)), the algorithm performed iterations to find the contours of the leaf (Fig. 2A(c)), after a maximum of 200 iterations a final mask was obtained (Fig. 2A(d)). An image of the leaf with contour was returned (Fig. 2B). The returned final mask is a matrix where pixels inside the mask were set to leaf temperature and pixels outside the mask were set to zero.

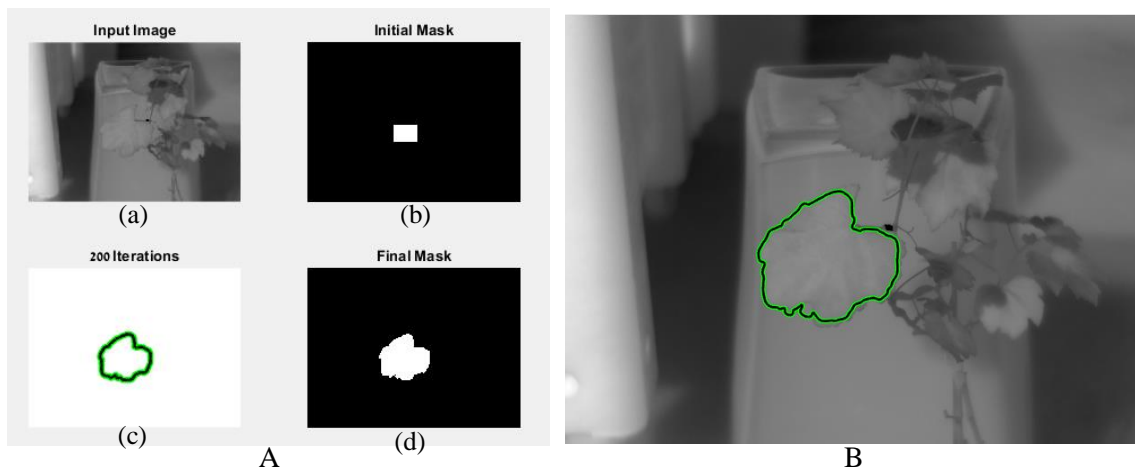


Figure 2. The output of the active contour algorithm. Includes image, initial and final mask (A) and image with final mask (B)

Feature selection

The feature selection was conducted by stepwise regression (Table 1) resulting in the four significant variables that were selected: MTD, MAD, percentile 25 and 98. There is a high correlation between the percentiles ($r=0.98$) based on the Pearson analysis. However, this result was expected due to the nature of the variable's construction (both related to temperature from the same leaf). Therefore, these 4 variables were selected to train the classification models.

Table 1. Stepwise regression- estimated coefficients, standard errors and p-value

Variable	Estimated coefficients	Standard errors	P-value
Perc-98	0.2569	0.0495	2.4678e-07
Perc-25	-0.2405	0.0497	1.5278e-06
MAD	-0.5341	0.1632	0.0011
MTD	0.0342	0.0183	0.0615
Perc-10	-0.0890	0.1093	0.4153
Perc-75	-0.1017	0.1351	0.4519
IQR	-0.0950	0.1362	0.4857
Tavg	-0.1058	0.2131	0.6196
Perc-2	-0.0273	0.0551	0.6204
Perc-90	0.0548	0.1402	0.6958
median	-0.0474	0.1315	0.7186
Tmin	-0.0154	0.0441	0.7275
Tmax	-0.0154	0.0441	0.7275
STD	-0.0528	0.3084	0.8640

Classification models

Each of the five models was optimized to determine the parameters for using the model. The optimal parameters obtained for each model were:

- Decision Tree: 20 maximum splits, Gini's diversity index split criterion
- Logistic Regression: Threshold 0.5
- Naive Bayes: Gaussian Kernel
- SVM: Gaussian Kernel using a scale of 2.6, box constraint level of 1
- Ensemble: Decision Tree learner type, Boosted Trees using LogitBoost method, 20 maximum splits, 30 learners using a learning rate of 0.1

Table 2 presents the accuracy, F1 score precision and recall of the constructed models using the different data sets as described above. All models performed rather well with small differences between the models. When considering the three sets, the highest results were achieved by the SVM model with small differences between the performance measures of the different sets.

To determine the earliest day for downy mildew detection, the accuracy was tested for images that were acquired 4 and 7 days after infection (Table 3). To avoid bias, these data sets were balanced. The results showed a difference of about 17% between the accuracy of the 4th day and the 7th day after infection. A possible reason why the accuracy results are not as high as expected is that the label whether the leaf was infected or healthy was set according to whether downy mildew application was applied or not (on leaves marked a priori) and not according to the actual development of the disease. The treatment dosages were different between the campaigns, which could have led to different levels of disease development. Therefore, this should be considered in further research.

Table 2. Results of all models on all data sets

Model	Performance measures [%]		0,1,2,4,7 days after infection		0,2,4,7 days after infection		0,4,7 days after infection	
Decision Tree	Accuracy	F1	65.8	70.4	68.5	72.6	69.3	64.1
	score		73.4	67.7	72.3	72.9	73.9	56.6
	Precision	Recall						
Logistic Regression	Accuracy	F1	60.6	71.0	62.9	69.7	65.4	62.2
	score		63.6	80.5	65.4	74.5	66.2	58.6
	Precision	Recall						
NB	Accuracy	F1	66.4	70.7	70.0	72.4	69.6	66.9
	score		74.3	67.3	76.5	68.7	70.1	63.3
	Precision	Recall						
SVM	Accuracy	F1	69.2	74.9	71.1	74.9	70.9	67.6
	score		73.5	76.3	74.6	75.2	73.8	62.3
	Precision	Recall						
Ensemble	Accuracy	F1	65.9	71.2	69.8	72.9	70.0	66.4
	score		72.3	70.2	75.1	70.8	72.7	61.1
	Precision	Recall						

Table 3. SVM classification results from balanced data sets

Model	Performance measures [%]		0,4 days after infection 256 infected, 256 healthy		0,7 days after infection 145 infected, 145 healthy	
SVM	Accuracy	F1 score	66.4	64.6	83.1	83.0
	Precision	Recall	68.3	61.3	83.3	82.8

In summary, the best model to classify healthy and infected leaves for all data was SVM, resulting in 69.2% accuracy and 74.9% F1 score in the cross-validation process. The accuracy was higher by 2.8% from the next best model which was Naïve Bayes and the F1 score was higher by 4.2%. The model used the following selected features MTD, MAD, percentile 25 and 98 which were selected by stepwise regression.

The highest accuracy was obtained from the set of 7 days after infection with an accuracy of 83.1% and an F1 score of 83%. However, also a specialist pathologist can detect that the leaf is infected 7 days after infection. The SVM model can detect infections also 4 days after infection with an accuracy of 66.4% and an F1 score of 64.6% when a specialist pathologist cannot detect that the leaf is infected.

Conclusions

The results indicate the feasibility of detecting infected leaves in thermograms, even before visible symptoms of downy mildew appear.

Ongoing research is focused on analyzing also actual disease severity. Additionally, due to the large difference between the results of days 4 and 7, new acquisitions will include also 5 and 6 days after infection to determine the earliest day for grapevine downy mildew detection and to expand the database.

Acknowledgements

Financial support for this research was provided by the Israel Innovation Authority, project no 67268. Partial support was provided by Ben-Gurion University of the Negev through the Rabbi W. Gunther Plaut Chair in Manufacturing Engineering and by the Molcho Center fund. We wish to thank Mr. S. Ozer for his support during the image acquisition campaigns. We are also grateful to Z. Halamish, N. Guedj, A. Himmel and A. Movchan for their excellent support on growing the plants and on conducting the experiments.

References

- Alchanatis, V., Cohen, Y., Cohen, S., Moller, M., Sprinstin, M., Meron, M., Tsipris, J., Saranga, Y., & Sela, E. (2010). Evaluation of different approaches for estimating and mapping crop water status in cotton with thermal imaging. *Precision Agriculture*, 11(1), 27–41. <https://doi.org/10.1007/s11119-009-9111-7>
- Anderson, M. C., hain, C., Otkin, J., Zhan, X., Mo, K., Svoboda, M., Wardlow, B., & Pimstein, A. (2013). An intercomparison of drought indicators based on thermal remote sensing and NLDAS-2 simulations with U.S. drought monitor classifications. *Journal of Hydrometeorology*, 14(4), 1035–1056. <https://doi.org/10.1175/JHM-D-12-0140.1>
- Ashourloo, D., Mobasheri, M. R., & Huete, A. (2014). Evaluating the effect of different wheat rust disease symptoms on vegetation indices using hyperspectral measurements. *Remote Sensing*, 6(6), 5107–5123. <https://doi.org/10.3390/rs6065107>
- Aziz, A., Poinssot, B., Daire, X., Adrian, M., Bézier, A., Lambert, B., Joubert, J. M., & Pugin, A. (2003). Laminarin Elicits Defense Responses in Grapevine and Induces Protection Against Botrytis cinerea and Plasmopara viticola. *Molecular Plant-Microbe Interactions*, 16(12), 1118–1128. <https://doi.org/10.1094/MPMI.2003.16.12.1118>
- Banerjee, K., Krishnan, P., & Mridha, N. (2018). Application of thermal imaging of wheat crop canopy to estimate leaf area index under different moisture stress conditions. *Biosystems Engineering*, 166(November), 13–27. <https://doi.org/10.1016/j.biosystemseng.2017.10.012>
- Barbagallo, S., Consoli, S., & Russo, A. (2009). A one-layer satellite surface energy balance for estimating evapotranspiration rates and crop water stress indexes. *Sensors*, 9(1), 1–21. <https://doi.org/10.3390/s90100001>
- Barbedo, J. G. A. (2016). A review on the main challenges in automatic plant disease identification based on visible range images. *Biosystems Engineering*, 144, 52–60. <https://doi.org/10.1016/j.biosystemseng.2016.01.017>
- Berdugo, C. A., Zito, R., Paulus, S., & Mahlein, A. K. (2014). Fusion of sensor data for the detection and differentiation of plant diseases in cucumber. *Plant Pathology*, 63(6), 1344–1356. <https://doi.org/10.1111/ppa.12219>
- Bian, J., Zhang, Z., Chen, J., Chen, H., Cui, C., Li, X., Chen, S., & Fu, Q. (2019). Simplified evaluation of cotton water stress using high resolution unmanned aerial vehicle thermal imagery. *Remote Sensing*, 11(3). <https://doi.org/10.3390/rs11030267>
- Bulanon, D. M., Burks, T. F., & Alchanatis, V. (2008). Study on temporal variation in citrus canopy using thermal imaging for citrus fruit detection. *Biosystems Engineering*, 101(2), 161–171. <https://doi.org/10.1016/j.biosystemseng.2008.08.002>
- Calderón, R., Montes-Borrego, M., Landa, B. B., Navas-Cortés, J. A., & Zarco-Tejada, P. J. (2014). Detection of downy mildew of opium poppy using high-resolution multi-spectral and thermal imagery acquired with an unmanned aerial vehicle. *Precision Agriculture*, 15(6), 639–661. <https://doi.org/10.1007/s11119-014-9360-y>
- Calderón, R., Navas-Cortés, J. A., Lucena, C., & Zarco-Tejada, P. J. (2013). High-resolution airborne hyperspectral and thermal imagery for early detection of Verticillium wilt of

- olive using fluorescence, temperature and narrow-band spectral indices. *Remote Sensing of Environment*, 139, 231–245. <https://doi.org/10.1016/j.rse.2013.07.031>
- Carlson, T. (2007). An overview of the “triangle method” for estimating surface evapotranspiration and soil moisture from satellite imagery. *Sensors*, 7(8), 1612–1629. <https://doi.org/10.3390/s7081612>
- Caro, S. G. (2014). Infection and spread of *Peronospora sparsa* on *Rosa* sp.(Berk.) - a microscopic and a thermographic approach. *Doctoral Dissertation, University of Bonn, Germany*.
- Cerutti, G., Tougne, L., Vacavant, A., Coquin, D., Cerutti, G., Tougne, L., Vacavant, A., Coquin, D., & Active, A. P. (2013). *A Parametric Active Polygon for Leaf Segmentation and Shape Estimation To cite this version : HAL Id : hal-00622269 A Parametric Active Polygon for Leaf Segmentation and Shape Estimation*.
- Chaerle, L., Hagenbeek, D., De Bruyne, E., Valcke, R., & Van Der Straeten, D. (2004). Thermal and chlorophyll-fluorescence imaging distinguish plant-pathogen interactions at an early stage. *Plant and Cell Physiology*, 45(7), 887–896. <https://doi.org/10.1093/pcp/pch097>
- Chan, T. F., & Vese, L. A. (2001). Active contours without edges. *IEEE Transactions on Image Processing*, 10(2), 266–277. <https://doi.org/10.1109/83.902291>
- Correia, M. J., Chaves, M. M. C., & pereira, J. S. (1990). Afternoon Depression In Photosynthesis in Grapevine Leaves—Evidence for a High Light Stress Effect. *Journal of Experimental Botany*, 41(4), 417–426. <https://doi.org/10.1093/jxb/41.4.417>
- Cui, S., Ling, P., Zhu, H., & Keener, H. M. (2018). Plant pest detection using an artificial nose system: A review. *Sensors (Switzerland)*, 18(2), 1–18. <https://doi.org/10.3390/s18020378>
- Dhaka, V. S., Meena, S. V., Rani, G., Sinwar, D., Kavita, Ijaz, M. F., & Woźniak, M. (2021). A survey of deep convolutional neural networks applied for prediction of plant leaf diseases. *Sensors*, 21(14), 4749. <https://doi.org/10.3390/s21144749>
- Farsund, A. A., Daugbjerg, C., & Langhelle, O. (2015). Food security and trade: reconciling discourses in the Food and Agriculture Organization and the World Trade Organization. *Food Security*, 7(2), 383–391. <https://doi.org/10.1007/s12571-015-0428-y>
- Fernandez-Gallego, J. A., Buchailot, M. L., Gutiérrez, N. A., Nieto-Taladriz, M. T., Araus, J. L., & Kefauver, S. C. (2019). Automatic wheat ear counting using thermal imagery. *Remote Sensing*, 11(7), 1–13. <https://doi.org/10.3390/rs11070751>
- Ford, T., Science, F., Eastern, M., Natural, M. M., & Eastern, M. (2019). *Remote Sensing in Precision Agriculture on Winter Wheat and a Subsurface Drip Irrigated (SDI) Corn Field 2019 ASABE Annual International Meeting*. 2–11.
- Frank, K. A. (2000). impact of a confounding variable on a regression coefficient. *Sociological Methods & Research*, 29(2), 147–194.
- Fuchs, M. (1990). Infrared measurement of canopy temperature and detection of plant water stress. *Theoretical and Applied Climatology*, 42(4), 253–261. <https://doi.org/10.1007/BF00865986>
- García-Tejero, I. F., Rubio, A. E., Viñuela, I., Hernández, A., Gutiérrez-Gordillo, S., Rodríguez-Pleguezuelo, C. R., & Durán-Zuazo, V. H. (2018). Thermal imaging at plant level to assess the crop-water status in almond trees (cv. Guara) under deficit irrigation strategies. *Agricultural Water Management*, 208(May 2017), 176–186. <https://doi.org/10.1016/j.agwat.2018.06.002>
- Gebbers, R., & Adamchuk, V. I. (2010). *Precision Agriculture and Food Security* (Vol. 327, Issue 5967). <https://doi.org/10.1126/science.1183899>
- Gessler, C., Pertot, I., & Perazzolli, M. (2011). Plasmopara viticola: A review of knowledge on downy mildew of grapevine and effective disease management. *Phytopathologia*

- Mediterranea*, 50(1), 3–44. https://doi.org/10.14601/Phytopathol_Mediterr-9360
- Golhani, K., Balasundram, S. K., Vadamalai, G., & Pradhan, B. (2018). A review of neural networks in plant disease detection using hyperspectral data. *Information Processing in Agriculture*, 5(3), 354–371. <https://doi.org/10.1016/j.inpa.2018.05.002>
- Gonzalez-Dugo, V., Zarco-Tejada, P., Nicolás, E., Nortes, P. A., Alarcón, J. J., Intrigliolo, D. S., & Fereres, E. (2013). Using high resolution UAV thermal imagery to assess the variability in the water status of five fruit tree species within a commercial orchard. *Precision Agriculture*, 14(6), 660–678. <https://doi.org/10.1007/s11119-013-9322-9>
- Grant, O. M., Chaves, M. M., & Jones, H. G. (2006). Optimizing thermal imaging as a technique for detecting stomatal closure induced by drought stress under greenhouse conditions. *Physiologia Plantarum*, 127(3), 507–518. <https://doi.org/10.1111/j.1399-3054.2006.00686.x>
- Granum, E., Pérez-Bueno, M. L., Calderón, C. E., Ramos, C., de Vicente, A., Cazorla, F. M., & Barón, M. (2015). Metabolic responses of avocado plants to stress induced by *Rosellinia necatrix* analysed by fluorescence and thermal imaging. *European Journal of Plant Pathology*, 142(3), 625–632. <https://doi.org/10.1007/s10658-015-0640-9>
- Gutiérrez, S., Diago, M. P., Fernández-Navales, J., & Tardaguila, J. (2018). Vineyard water status assessment using on-the-go thermal imaging and machine learning. *PLoS ONE*, 13(2), 1–19. <https://doi.org/10.1371/journal.pone.0192037>
- Hassan-Esfahani, L., Torres-Rua, A., Jensen, A., & McKee, M. (2015). Assessment of surface soil moisture using high-resolution multi-spectral imagery and artificial neural networks. *Remote Sensing*, 7(3), 2627–2646. <https://doi.org/10.3390/rs70302627>
- Jones, H. G. (1999). Use of infrared thermometry for estimation of stomatal conductance as a possible aid to irrigation scheduling. *Agricultural and Forest Meteorology*, 95(3), 139–149. [https://doi.org/10.1016/S0168-1923\(99\)00030-1](https://doi.org/10.1016/S0168-1923(99)00030-1)
- Khanal, S., Fulton, J., & Shearer, S. (2017). An overview of current and potential applications of thermal remote sensing in precision agriculture. *Computers and Electronics in Agriculture*, 139, 22–32. <https://doi.org/10.1016/j.compag.2017.05.001>
- Kiefer, B., Riemann, M., Büche, C., Kassemeyer, H. H., & Nick, P. (2002). The host guides morphogenesis and stomatal targeting in the grapevine pathogen *Plasmopara viticola*. *Planta*, 215(3), 387–393. <https://doi.org/10.1007/s00425-002-0760-2>
- Lindenthal, M., Steiner, U., Dehne, H. W., & Oerke, E. C. (2005). Effect of downy mildew development on transpiration of cucumber leaves visualized by digital infrared thermography. *Phytopathology*, 95(3), 233–240. <https://doi.org/10.1094/PHYTO-95-0233>
- Linke, M., Beuche, H., Geyer, M., & Hellebrand, H. J. (2000). Possibilities and limits of the use of thermography for the examination of horticultural products. *Agrartechnische Forschung*, 6(2), 110–114. <https://www.cabdirect.org/cabdirect/abstract/20013076270>
- Mahlein, A.-K. (2016). Plant disease detection by imaging sensors- parallels and specific demands for precision agriculture and plant phenotyping. *Plant Disease*, 100(2), 1–11. <https://doi.org/10.1007/s13398-014-0173-7.2>
- Mahlein, A. K., Oerke, E. C., Steiner, U., & Dehne, H. W. (2012). Recent advances in sensing plant diseases for precision crop protection. *European Journal of Plant Pathology*, 133(1), 197–209. <https://doi.org/10.1007/s10658-011-9878-z>
- Martinelli, F., Scalenghe, R., Davino, S., Panno, S., Scuderi, G., Ruisi, P., Villa, P., Stroppiana, D., Boschetti, M., Goulart, L. R., Davis, C. E., & Dandekar, A. M. (2015). Advanced methods of plant disease detection. A review. *Agronomy for Sustainable Development*, 35(1), 1–25. <https://doi.org/10.1007/s13593-014-0246-1>
- O’Shaughnessy, S. A., & Evett, S. R. (2010). Canopy temperature based system effectively schedules and controls center pivot irrigation of cotton. *Agricultural Water Management*,

- 97(9), 1310–1316. <https://doi.org/10.1016/j.agwat.2010.03.012>
- O'Shaughnessy, Susan A., Evett, S. R., Colaizzi, P. D., & Howell, T. A. (2012). A crop water stress index and time threshold for automatic irrigation scheduling of grain sorghum. *Agricultural Water Management*, 107, 122–132. <https://doi.org/10.1016/j.agwat.2012.01.018>
- Oberti, R., Marchi, M., Tirelli, P., Calcante, A., Iriti, M., & Borghese, A. N. (2014). Automatic detection of powdery mildew on grapevine leaves by image analysis: Optimal view-angle range to increase the sensitivity. *Computers and Electronics in Agriculture*, 104, 1–8. <https://doi.org/10.1016/j.compag.2014.03.001>
- Oerke, E. C., Fröhling, P., & Steiner, U. (2011). Thermographic assessment of scab disease on apple leaves. *Precision Agriculture*, 12(5), 699–715. <https://doi.org/10.1007/s11119-010-9212-3>
- Oerke, E. C., Steiner, U., Dehne, H. W., & Lindenthal, M. (2006). Thermal imaging of cucumber leaves affected by downy mildew and environmental conditions. *Journal of Experimental Botany*, 57(9), 2121–2132. <https://doi.org/10.1093/jxb/erj170>
- Osroosh, Y., Khot, L. R., & Peters, R. T. (2018). Economical thermal-RGB imaging system for monitoring agricultural crops. *Computers and Electronics in Agriculture*, 147(August 2017), 34–43. <https://doi.org/10.1016/j.compag.2018.02.018>
- Osroosh, Y., Troy Peters, R., Campbell, C. S., & Zhang, Q. (2015). Automatic irrigation scheduling of apple trees using theoretical crop water stress index with an innovative dynamic threshold. *Computers and Electronics in Agriculture*. <https://doi.org/10.1016/j.compag.2015.09.006>
- Pedersen, S. M., & Lind, K. M. (2017). *Precision Agriculture : Technology and Economic*. <https://doi.org/https://doi-org.ezproxy.bgu.ac.il/10.1007/978-3-319-68715-5>
- Ray, D. K., Mueller, N. D., West, P. C., & Foley, J. A. (2013). Yield Trends Are Insufficient to Double Global Crop Production by 2050. *PLoS ONE*, 8(6). <https://doi.org/10.1371/journal.pone.0066428>
- Sankaran, S., Mishra, A., Ehsani, R., & Davis, C. (2010). A review of advanced techniques for detecting plant diseases. *Computers and Electronics in Agriculture*, 72(1), 1–13. <https://doi.org/10.1016/j.compag.2010.02.007>
- Savary, S., Ficke, A., Aubertot, J. N., & Hollier, C. (2012). Crop losses due to diseases and their implications for global food production losses and food security. *Food Security*, 4(4), 519–537. <https://doi.org/10.1007/s12571-012-0200-5>
- Seelan, S. K., Laguet, S., Casady, G. M., & Seielstad, G. A. (2003). *Remote sensing applications for precision agriculture : A learning community approach*. 88, 157–169. <https://doi.org/10.1016/j.rse.2003.04.007>
- Shafian, S., & Maas, S. J. (2015). Index of soil moisture using raw Landsat image digital count data in Texas High Plains. *Remote Sensing*, 7(3), 2352–2372. <https://doi.org/10.3390/rs70302352>
- Soliman, A., Heck, R. J., Brenning, A., Brown, R., & Miller, S. (2013). Remote sensing of soil moisture in vineyards using airborne and ground-based thermal inertia data. *Remote Sensing*, 5(8), 3729–3748. <https://doi.org/10.3390/rs5083729>
- Stajanko, D., Lakota, M., & Hočevár, M. (2004). Estimation of number and diameter of apple fruits in an orchard during the growing season by thermal imaging. *Computers and Electronics in Agriculture*, 42(1), 31–42. [https://doi.org/10.1016/S0168-1699\(03\)00086-3](https://doi.org/10.1016/S0168-1699(03)00086-3)
- Stoll, M., Schultz, H. R., Baecker, G., & Berkelmann-Loehnertz, B. (2008). Early pathogen detection under different water status and the assessment of spray application in vineyards through the use of thermal imagery. *Precision Agriculture*, 9(6), 407–417. <https://doi.org/10.1007/s11119-008-9084-y>

- Sullivan, D. G., Shaw, J. N., Mask, P. L., Rickman, D., Guertal, E. A., Luvall, J., & Wersinger, J. M. (2004). Evaluation of multispectral data for rapid assessment of wheat straw residue cover. *Soil Science Society of America Journal*, 68(6), 2007–2013.
- Testi, L., Goldhamer, D. A., Iniesta, F., & Salinas, M. (2008). Crop water stress index is a sensitive water stress indicator in pistachio trees. *Irrigation Science*, 26(5), 395–405. <https://doi.org/10.1007/s00271-008-0104-5>
- Vollmer, M., & Möllmann, K.-P. (2017). Fundamentals of Infrared Thermal Imaging. In *Infrared Thermal Imaging*. <https://doi.org/10.1002/9783527693306.ch1>
- Wen, D. M., Chen, M. X., Zhao, L., Ji, T., Li, M., & Yang, X. T. (2019). Use of thermal imaging and Fourier transform infrared spectroscopy for the pre-symptomatic detection of cucumber downy mildew. *European Journal of Plant Pathology*, 155(2), 405–416. <https://doi.org/10.1007/s10658-019-01775-2>
- Zhou, J., Khot, L. R., Boydston, R. A., Miklas, P. N., & Porter, L. (2018). Low altitude remote sensing technologies for crop stress monitoring: a case study on spatial and temporal monitoring of irrigated pinto bean. *Precision Agriculture*, 19(3), 555–569. <https://doi.org/10.1007/s11119-017-9539-0>

Appendix C - Experiment map and schedule

Exp. Number	Start date	Fungi Application & Incubation O. N	TP1	TP2	TP3	TP4	TP5	Number of plants
9848	24/12/2019	29/12/2019	26/12/2019	30/12/2019 + healthy	31/12/2019	02/01/2020	05/01/2020	34+2
10081	07/01/2020	12/01/2020	12/01/2020	13/01/2020	14/01/2020	16/01/2020	19/01/2020	30
10082	14/01/2020	19/01/2020	19/01/2020	20/01/2020	21/01/2020	23/01/2020	26/01/2020	27
10085	25/02/2020	01/03/2020	25/02/2020	02/03/2020	03/03/2020 + healthy	05/03/2020	08/03/2020	27+7
10088	24/03/2020	29/03/2020	26/03/2020	30/03/2020	31/03/2020	02/04/2020	05/04/2020	27
10446	15/10/2020	21/10/2020	-	25/10/2020	26/10/2020	27/10/2020	28/10/2020	15

Green cells- executed imaging days

Blue cells- canceled days due to weather conditions

Red cells- canceled days by Evogene because of technical conditions

Appendix D - Data sets

Link: <https://github.com/BarCohenBGU/database.git>

'All data new'- the full database included 1403 records

'normalized_by_severity'- included 1097 records with actual disease severity of 5 or higher

'features_without_outliers'- included 1012 records after removing outlier

Appendix E - Daily data set

Link: <https://github.com/BarCohenBGU/daily-dataset.git>

Name of file: 'Data_27_10_2020_daily'

Appendix F – Normalization and Standardization

The accuracy of the SVM model with different methods of normalization.

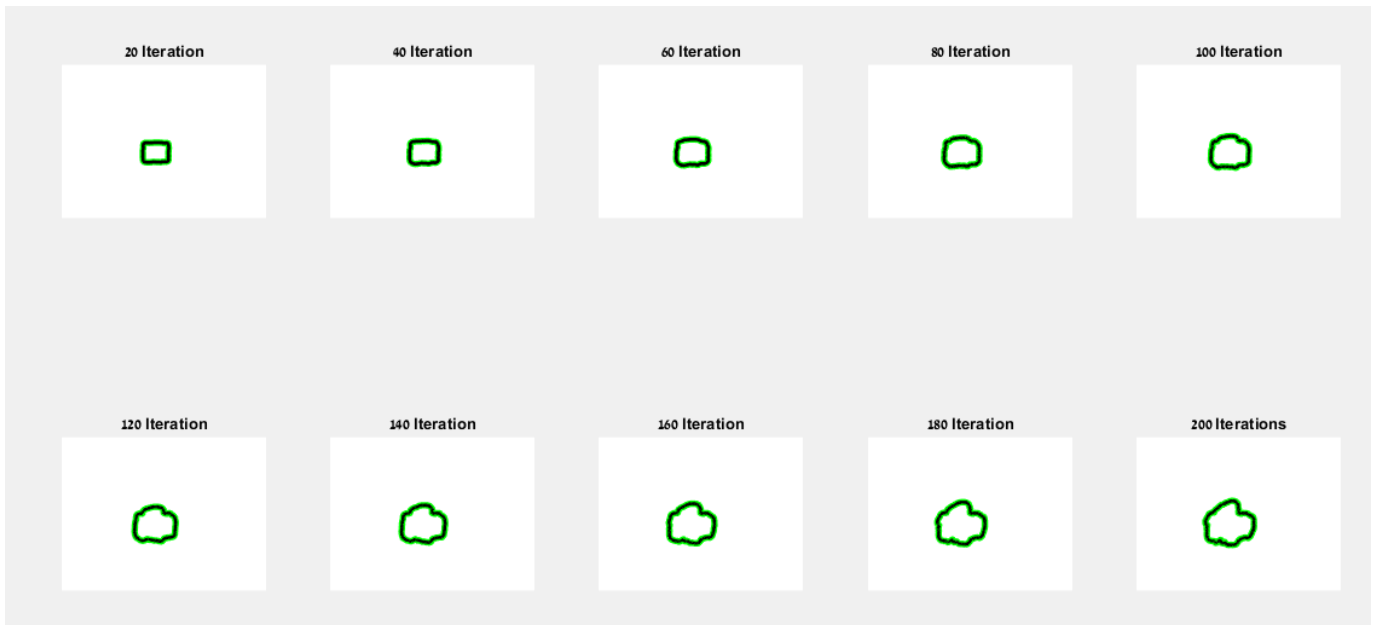
Dataset	Z-score normalization	Min–Max normalization
All data	0.8103	0.8172
Balanced data	0.805	0.7914

Appendix G – Active contour algorithm

Link: <https://github.com/BarCohenBGU/active-contour-algorithm.git>

Included functions: main, active contours, get curvature, minimize energy, re initialization, read points, xlscol

Appendix H – Active contour iterations

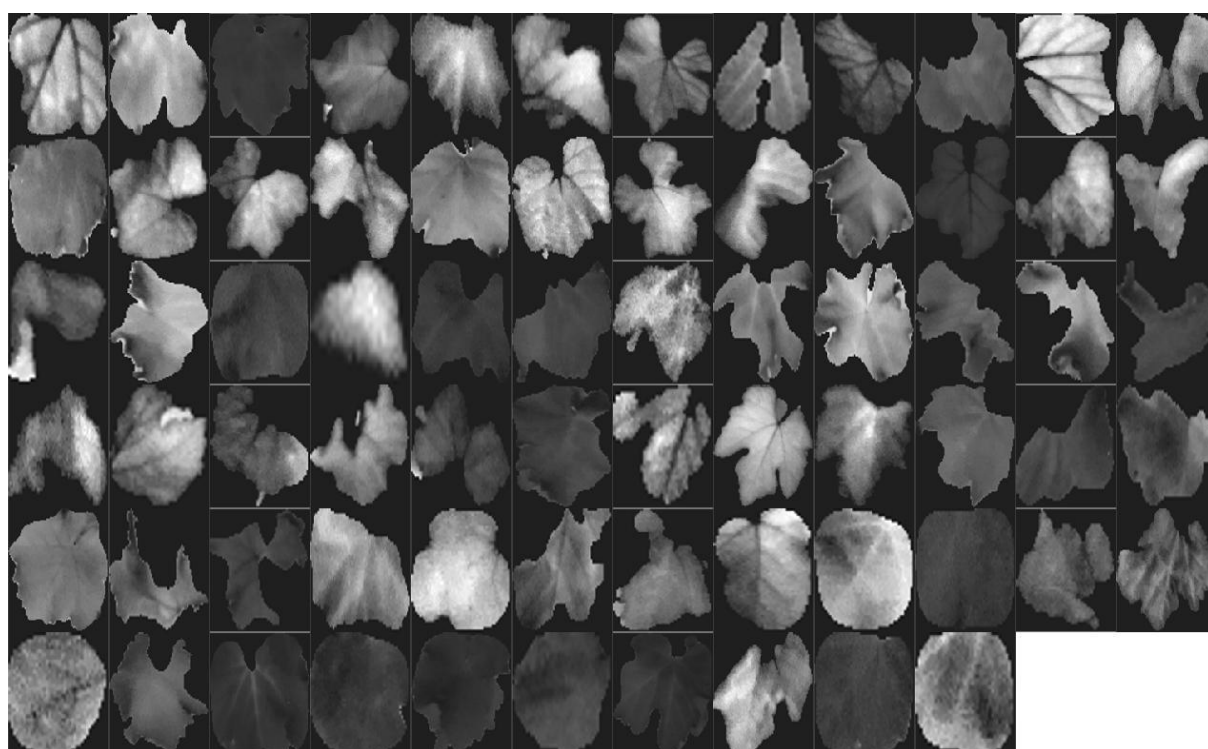


Appendix I – Images after edge detection algorithm

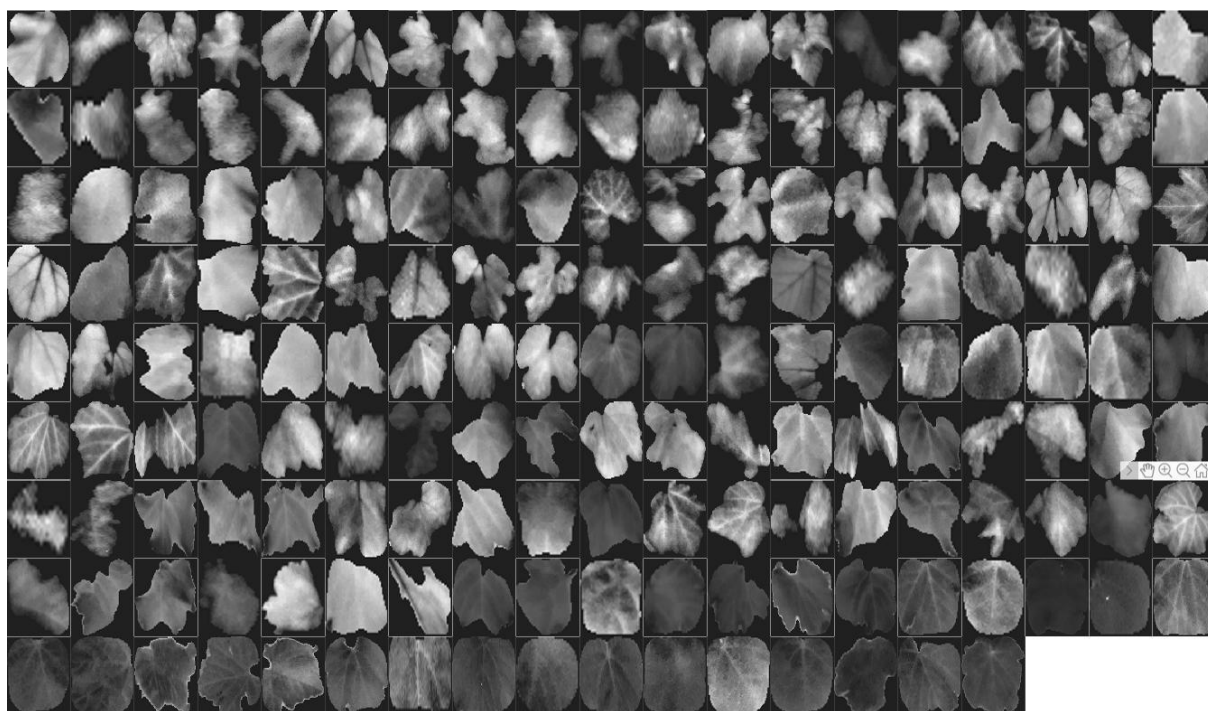
Healthy



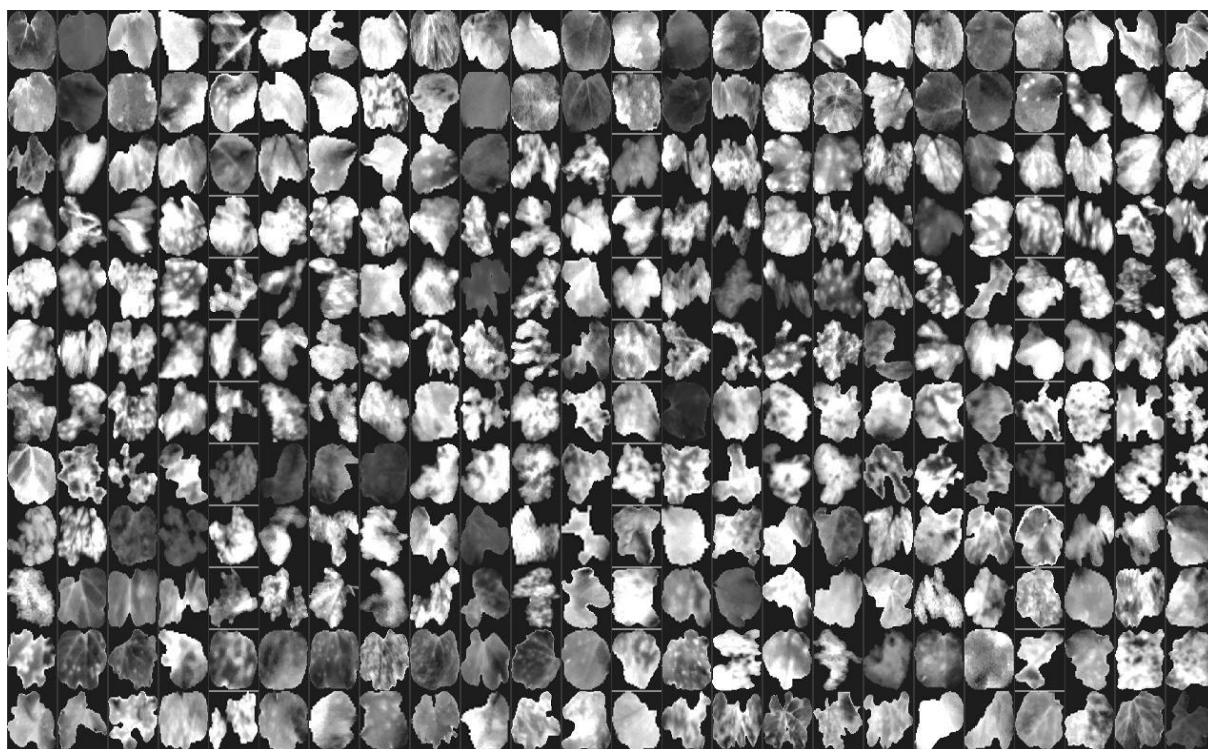
1 day after infection



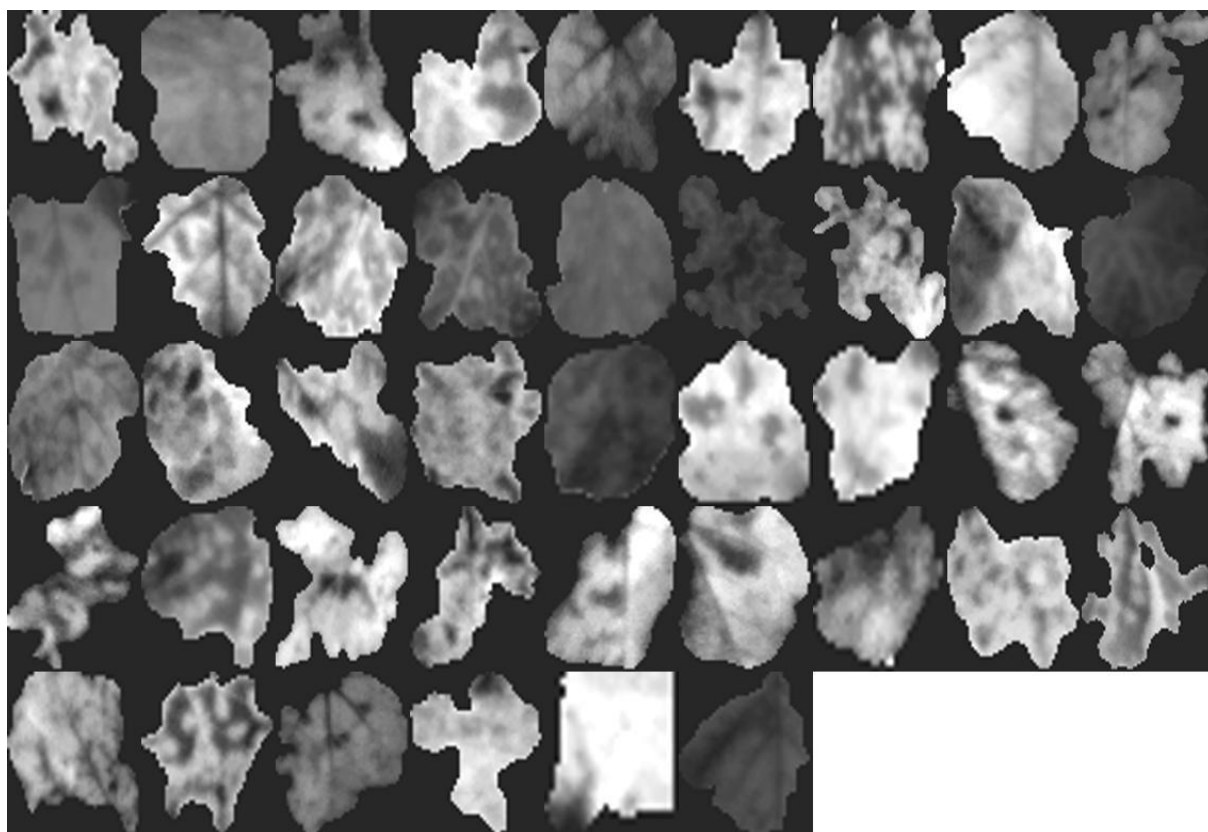
2 days after infection



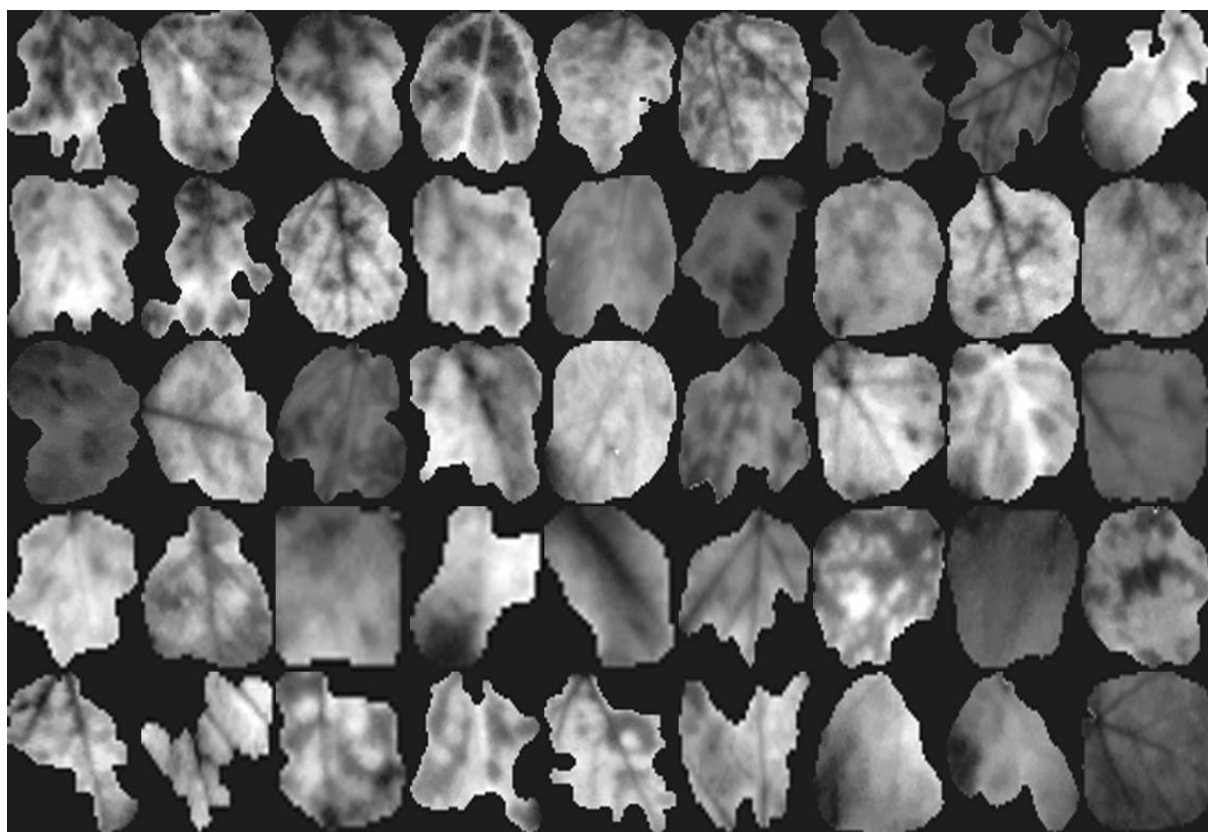
4 days after infection



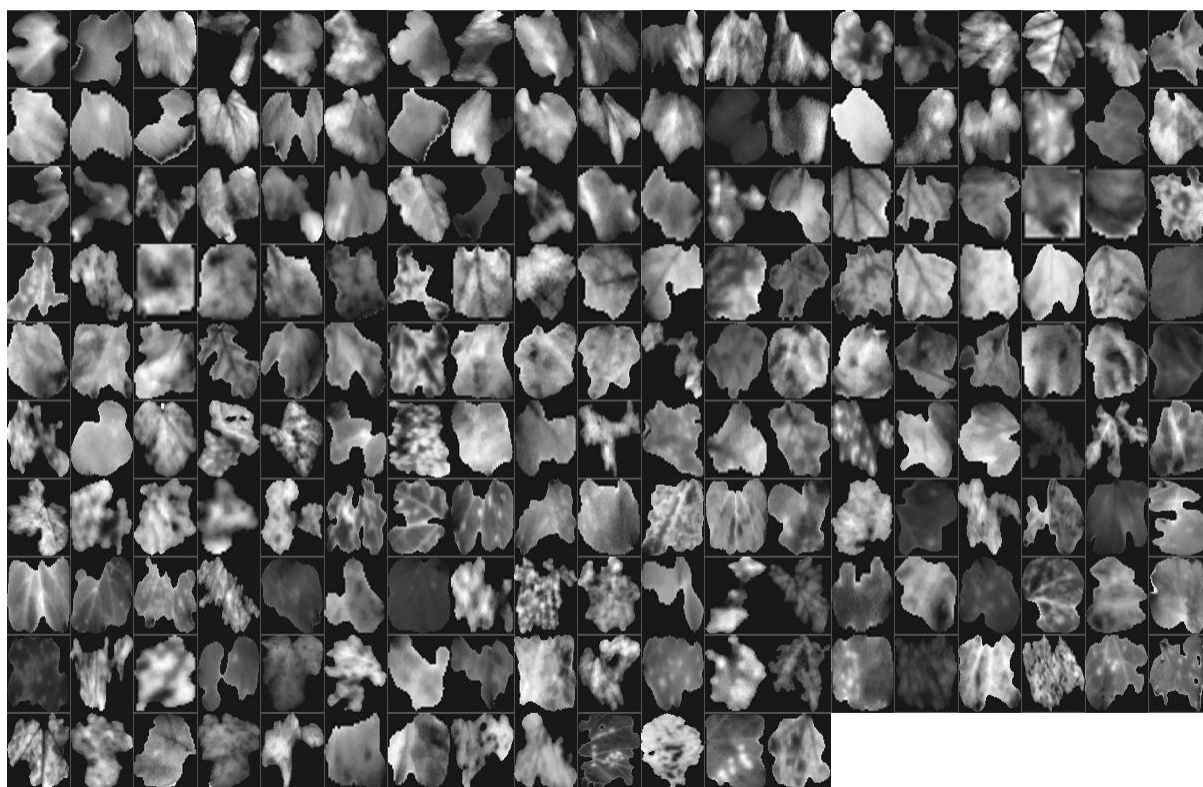
5 days after infection



6 days after infection



7 days after infection



Appendix J – Models

Link: <https://github.com/BarCohenBGU/classification-models.git>

Included functions: classification, classification daily, DecisionTree, NB, SVM, LogisticRegression, Ensemble

Appendix K – Statistics

Included Histograms and boxplots, Outliers, Correlation

Link: <https://github.com/BarCohenBGU/statistics.git>

Name of file: 'statistics'

Appendix L - Stepwise logistic regression

Stepwise with method "both":

```
Call:
glm(formula = Y ~ MTD + STD + Cv + Tavg.Tair + median.Tair +
    perc90.Tair + CWSI2, family = binomial, data = df)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.1545  -0.9113  -0.5136   0.9428   2.6276

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -4.1055     0.3610  -11.372  < 2e-16 ***
MTD             0.8840     0.1921   4.602 4.18e-06 ***
STD          -10.6517     2.1472  -4.961 7.02e-07 ***
Cv            69.9300    25.3766   2.756 0.005857 **
Tavg.Tair     -9.5336     2.4270  -3.928 8.56e-05 ***
median.Tair    3.6965     1.3432   2.752 0.005925 **
perc90.Tair    6.0920     1.2209   4.990 6.04e-07 ***
CWSI2          1.7728     0.4690   3.780 0.000157 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1386.2  on 1011  degrees of freedom
Residual deviance: 1131.2  on 1004  degrees of freedom
AIC: 1147.2

Number of Fisher Scoring iterations: 4
```

After removing median and Tavg

```
Call:
glm(formula = Y ~ MTD + STD + Cv + perc90.Tair + CWSI2, family = binomial,
    data = df)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.9813  -0.9106  -0.5501   0.9578   2.6795

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -4.09098     0.35413  -11.552  < 2e-16 ***
MTD             0.65434     0.17816   3.673 0.000240 ***
STD          -2.43727     1.13229  -2.153 0.031357 *
Cv            79.22257    24.76411   3.199 0.001379 **
perc90.Tair    0.27087     0.03883   6.975 3.06e-12 ***
CWSI2          1.64053     0.45752   3.586 0.000336 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1386.2  on 1011  degrees of freedom
Residual deviance: 1158.7  on 1006  degrees of freedom
AIC: 1170.7

Number of Fisher Scoring iterations: 4
```

Code:

Link: <https://github.com/BarCohenBGU/feature-selection.git>

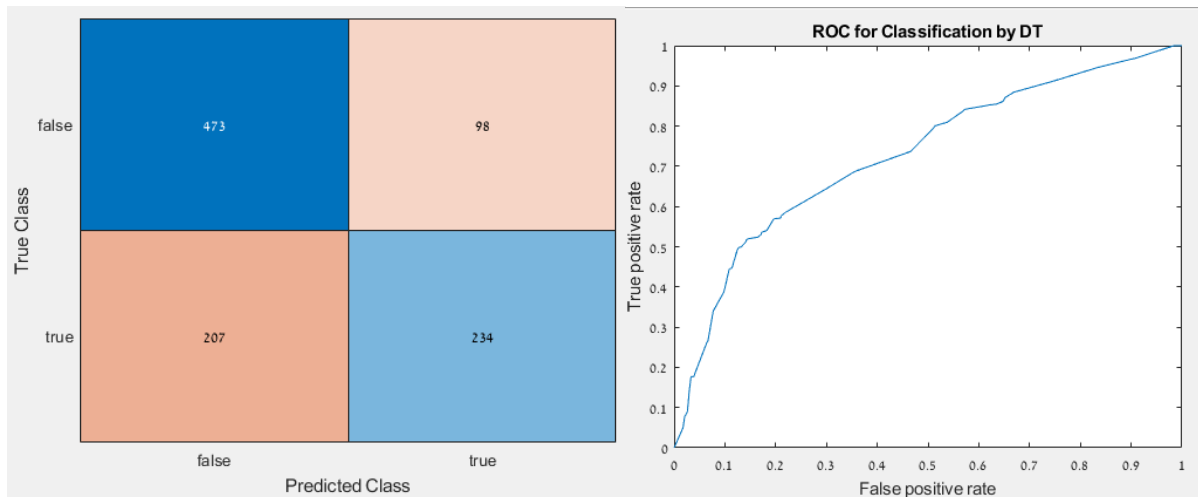
Included function: grapevine

Appendix M – Confusion matrixes and ROC curves

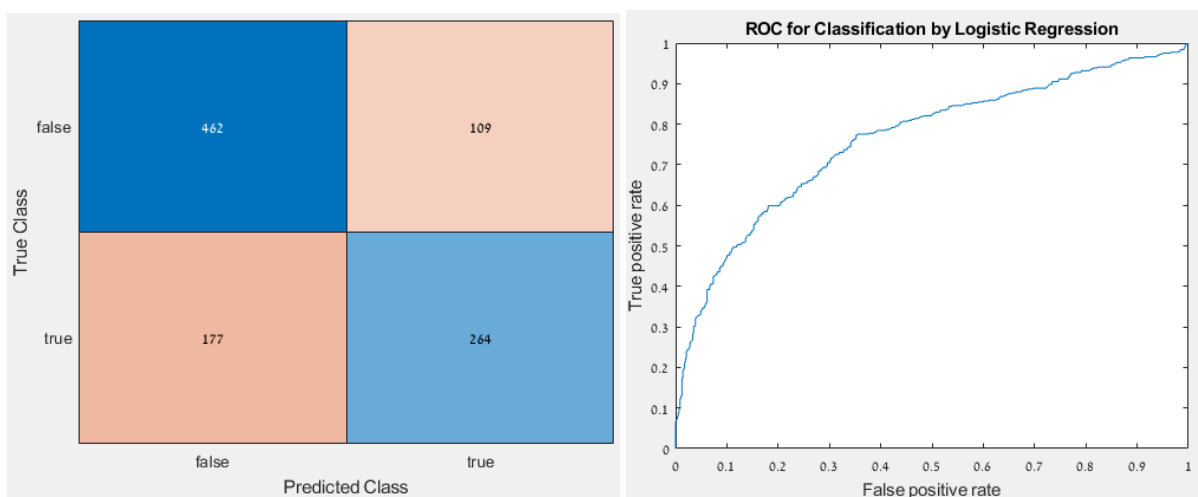
True= infected

False= healthy

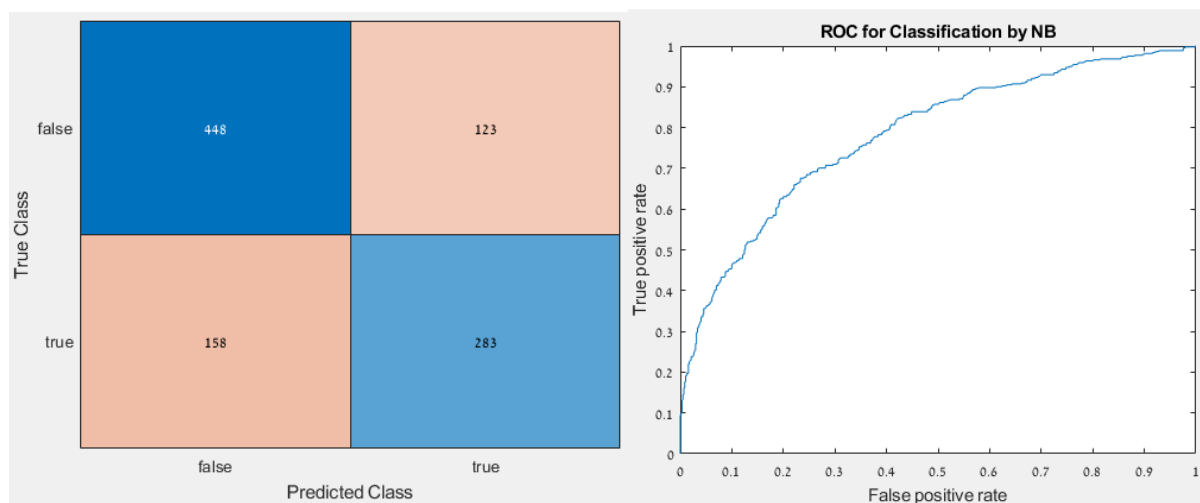
Decision tree



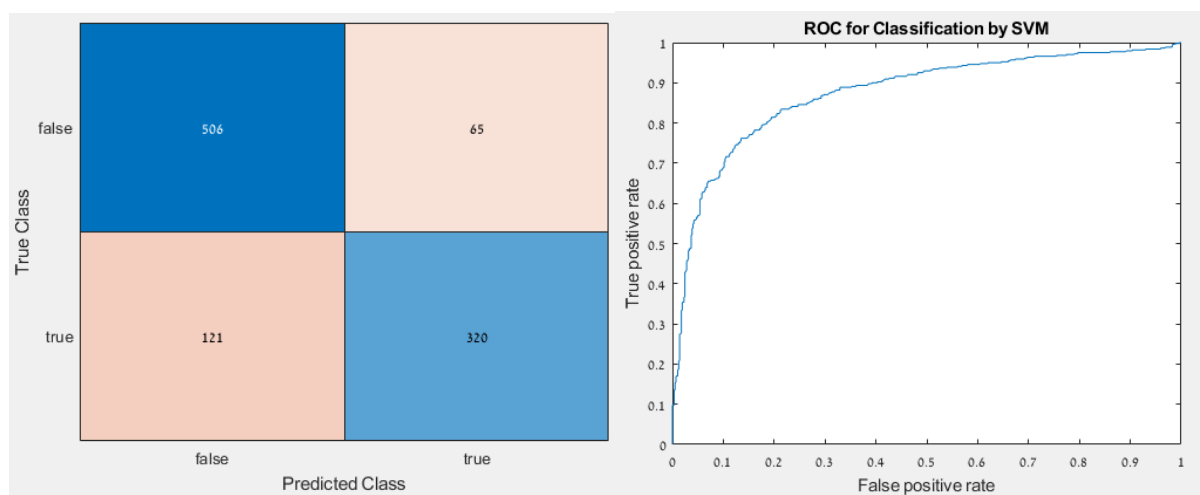
Logistic Regression



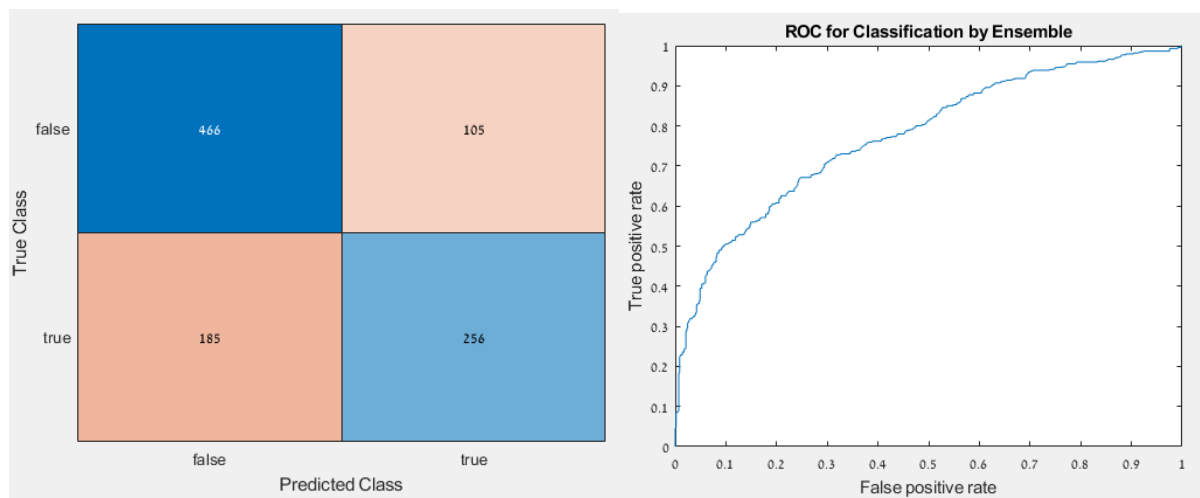
Naive Base



SVM



Ensemble



Appendix N – F-test and T-test for classification

f1 score		
NB	SVM	fold
0.686391	0.756098	1
0.651934	0.768293	2
0.629834	0.807229	3
0.725	0.760736	4
0.653846	0.781065	5
0.669401	0.774684	AVG

F-test

F-Test Two-Sample for Variances		
Variable 2 - F1 score, SVM	Variable 1- F1 score, NB	
0.774684089	0.669400928	Mean
0.000420184	0.00137312	Variance
5	5	Observations
4	4	df
	3.267902339	F
	0.138972663	P(F<=f) one-tail
	6.388232909	F Critical one-tail

t-test

t-Test: Two-Sample Assuming Equal Variances		
Variable 2 - F1 score, SVM	Variable 1 - F1 score, NB	
0.774684089	0.669400928	Mean
0.000420184	0.00137312	Variance
5	5	Observations
	0.000896652	Pooled Variance
	0	Hypothesized Mean Difference
	8	df
	-5.559259688	t Stat
	0.000267567	P(T<=t) one-tail
	1.859548038	t Critical one-tail
	0.000535133	P(T<=t) two-tail
	2.306004135	t Critical two-tail

Appendix O – F-test and T-test for best time acquisition

accuracy		
round 3	round 1	fold
0.882353	0.882353	1
0.833333	0.611111	2
0.777778	0.777778	3
0.777778	0.777778	4
0.764706	0.705882	5
0.80719	0.75098	AVG

F-test

F-Test Two-Sample for Variances		
Variable 2- accuracy, round 3	Variable 1 - accuracy, round 1	
0.807189542	0.750980392	Mean
0.002467	0.010073049	Variance
5	5	Observations
4	4	df
	4.083116883	F
	0.100879822	P(F<=f) one-tail
	6.388232909	F Critical one-tail

t-test

t-Test: Two-Sample Assuming Equal Variances		
Variable 2- accuracy, round 3	Variable 1 - accuracy, round 1	
0.807189542	0.750980392	Mean
0.002467	0.010073049	Variance
5	5	Observations
	0.006270024	Pooled Variance
	0	Hypothesized Mean Difference
	8	df
	-1.12238644	t Stat
	0.147125986	P(T<=t) one-tail
	1.859548038	t Critical one-tail
	0.294251973	P(T<=t) two-tail
	2.306004135	t Critical two-tail

References

Written References

- Alchanatis, V., Cohen, Y., Cohen, S., Moller, M., Sprinstin, M., Meron, M., Tsipris, J., Saranga, Y., & Sela, E. (2010). Evaluation of different approaches for estimating and mapping crop water status in cotton with thermal imaging. *Precision Agriculture*, 11(1), 27–41. <https://doi.org/10.1007/s11119-009-9111-7>
- Anderson, M. C., hain, C., Otkin, J., Zhan, X., Mo, K., Svoboda, M., Wardlow, B., & Pimstein, A. (2013). An intercomparison of drought indicators based on thermal remote sensing and NLDAS-2 simulations with U.S. drought monitor classifications. *Journal of Hydrometeorology*, 14(4), 1035–1056. <https://doi.org/10.1175/JHM-D-12-0140.1>
- Ashourloo, D., Mobasheri, M. R., & Huete, A. (2014). Evaluating the effect of different wheat rust disease symptoms on vegetation indices using hyperspectral measurements. *Remote Sensing*, 6(6), 5107–5123. <https://doi.org/10.3390/rs6065107>
- Aziz, A., Poinssot, B., Daire, X., Adrian, M., Bézier, A., Lambert, B., Joubert, J. M., & Pugin, A. (2003). Laminarin Elicits Defense Responses in Grapevine and Induces Protection Against Botrytis cinerea and Plasmopara viticola. *Molecular Plant-Microbe Interactions*, 16(12), 1118–1128. <https://doi.org/10.1094/MPMI.2003.16.12.1118>
- Banerjee, K., Krishnan, P., & Mridha, N. (2018). Application of thermal imaging of wheat crop canopy to estimate leaf area index under different moisture stress conditions. *Biosystems Engineering*, 166(November), 13–27. <https://doi.org/10.1016/j.biosystemseng.2017.10.012>
- Barbagallo, S., Consoli, S., & Russo, A. (2009). A one-layer satellite surface energy balance for estimating evapotranspiration rates and crop water stress indexes. *Sensors*, 9(1), 1–21. <https://doi.org/10.3390/s90100001>
- Barbedo, J. G. A. (2016). A review on the main challenges in automatic plant disease identification based on visible range images. *Biosystems Engineering*, 144, 52–60. <https://doi.org/10.1016/j.biosystemseng.2016.01.017>
- Berdugo, C. A., Zito, R., Paulus, S., & Mahlein, A. K. (2014). Fusion of sensor data for the detection and differentiation of plant diseases in cucumber. *Plant Pathology*, 63(6), 1344–1356. <https://doi.org/10.1111/ppa.12219>
- Bian, J., Zhang, Z., Chen, J., Chen, H., Cui, C., Li, X., Chen, S., & Fu, Q. (2019). Simplified evaluation of cotton water stress using high resolution unmanned aerial vehicle thermal imagery. *Remote Sensing*, 11(3). <https://doi.org/10.3390/rs11030267>
- Bulanon, D. M., Burks, T. F., & Alchanatis, V. (2008). Study on temporal variation in citrus canopy using thermal imaging for citrus fruit detection. *Biosystems Engineering*, 101(2), 161–171. <https://doi.org/10.1016/j.biosystemseng.2008.08.002>
- Calderón, R., Montes-Borrego, M., Landa, B. B., Navas-Cortés, J. A., & Zarco-Tejada, P. J. (2014). Detection of downy mildew of opium poppy using high-resolution multi-spectral and thermal imagery acquired with an unmanned aerial vehicle. *Precision Agriculture*, 15(6), 639–661. <https://doi.org/10.1007/s11119-014-9360-y>

- Calderón, R., Navas-Cortés, J. A., Lucena, C., & Zarco-Tejada, P. J. (2013). High-resolution airborne hyperspectral and thermal imagery for early detection of *Verticillium* wilt of olive using fluorescence, temperature and narrow-band spectral indices. *Remote Sensing of Environment*, 139, 231–245. <https://doi.org/10.1016/j.rse.2013.07.031>
- Carlson, T. (2007). An overview of the “triangle method” for estimating surface evapotranspiration and soil moisture from satellite imagery. *Sensors*, 7(8), 1612–1629. <https://doi.org/10.3390/s7081612>
- Caro, S. G. (2014). Infection and spread of *Peronospora sparsa* on *Rosa* sp.(Berk.) - a microscopic and a thermographic approach. *Doctoral Dissertation, University of Bonn, Germany*.
- Cerutti, G., Tougne, L., Vacavant, A., Coquin, D., Cerutti, G., Tougne, L., Vacavant, A., Coquin, D., & Active, A. P. (2013). *A Parametric Active Polygon for Leaf Segmentation and Shape Estimation To cite this version : HAL Id : hal-00622269 A Parametric Active Polygon for Leaf Segmentation and Shape Estimation*.
- Chaerle, L., Hagenbeek, D., De Bruyne, E., Valcke, R., & Van Der Straeten, D. (2004). Thermal and chlorophyll-fluorescence imaging distinguish plant-pathogen interactions at an early stage. *Plant and Cell Physiology*, 45(7), 887–896. <https://doi.org/10.1093/pcp/pch097>
- Chan, T. F., & Vese, L. A. (2001). Active contours without edges. *IEEE Transactions on Image Processing*, 10(2), 266–277. <https://doi.org/10.1109/83.902291>
- Correia, M. J., Chaves, M. M. C., & pereira, J. S. (1990). Afternoon Depression In Photosynthesis in Grapevine Leaves—Evidence for a High Light Stress Effect. *Journal of Experimental Botany*, 41(4), 417–426. <https://doi.org/10.1093/jxb/41.4.417>
- Cui, S., Ling, P., Zhu, H., & Keener, H. M. (2018). Plant pest detection using an artificial nose system: A review. *Sensors (Switzerland)*, 18(2), 1–18. <https://doi.org/10.3390/s18020378>
- Dhaka, V. S., Meena, S. V., Rani, G., Sinwar, D., Kavita, Ijaz, M. F., & Woźniak, M. (2021). A survey of deep convolutional neural networks applied for prediction of plant leaf diseases. *Sensors*, 21(14), 4749. <https://doi.org/10.3390/s21144749>
- Farsund, A. A., Daugbjerg, C., & Langhelle, O. (2015). Food security and trade: reconciling discourses in the Food and Agriculture Organization and the World Trade Organization. *Food Security*, 7(2), 383–391. <https://doi.org/10.1007/s12571-015-0428-y>
- Fernandez-Gallego, J. A., Buchaillet, M. L., Gutiérrez, N. A., Nieto-Taladriz, M. T., Araus, J. L., & Kefauver, S. C. (2019). Automatic wheat ear counting using thermal imagery. *Remote Sensing*, 11(7), 1–13. <https://doi.org/10.3390/rs11070751>
- Ford, T., Science, F., Eastern, M., Natural, M. M., & Eastern, M. (2019). *Remote Sensing in Precision Agriculture on Winter Wheat and a Subsurface Drip Irrigated (SDI) Corn Field 2019 ASABE Annual International Meeting*. 2–11.
- Frank, K. A. (2000). impact of a confounding variable on a regression coefficient. *Sociological Methods & Research*, 29(2), 147–194.

- Fuchs, M. (1990). Infrared measurement of canopy temperature and detection of plant water stress. *Theoretical and Applied Climatology*, 42(4), 253–261. <https://doi.org/10.1007/BF00865986>
- García-Tejero, I. F., Rubio, A. E., Viñuela, I., Hernández, A., Gutiérrez-Gordillo, S., Rodríguez-Pleguezuelo, C. R., & Durán-Zuazo, V. H. (2018). Thermal imaging at plant level to assess the crop-water status in almond trees (cv. Guara) under deficit irrigation strategies. *Agricultural Water Management*, 208(May 2017), 176–186. <https://doi.org/10.1016/j.agwat.2018.06.002>
- Gebbers, R., & Adamchuk, V. I. (2010). *Precision Agriculture and Food Security* (Vol. 327, Issue 5967). <https://doi.org/10.1126/science.1183899>
- Gessler, C., Pertot, I., & Perazzolli, M. (2011). Plasmopara viticola: A review of knowledge on downy mildew of grapevine and effective disease management. *Phytopathologia Mediterranea*, 50(1), 3–44. https://doi.org/10.14601/Phytopathol_Mediterr-9360
- Golhani, K., Balasundram, S. K., Vadamalai, G., & Pradhan, B. (2018). A review of neural networks in plant disease detection using hyperspectral data. *Information Processing in Agriculture*, 5(3), 354–371. <https://doi.org/10.1016/j.inpa.2018.05.002>
- Gonzalez-Dugo, V., Zarco-Tejada, P., Nicolás, E., Nortes, P. A., Alarcón, J. J., Intrigliolo, D. S., & Fereres, E. (2013). Using high resolution UAV thermal imagery to assess the variability in the water status of five fruit tree species within a commercial orchard. *Precision Agriculture*, 14(6), 660–678. <https://doi.org/10.1007/s11119-013-9322-9>
- Grant, O. M., Chaves, M. M., & Jones, H. G. (2006). Optimizing thermal imaging as a technique for detecting stomatal closure induced by drought stress under greenhouse conditions. *Physiologia Plantarum*, 127(3), 507–518. <https://doi.org/10.1111/j.1399-3054.2006.00686.x>
- Granum, E., Pérez-Bueno, M. L., Calderón, C. E., Ramos, C., de Vicente, A., Cazorla, F. M., & Barón, M. (2015). Metabolic responses of avocado plants to stress induced by Rosellinia necatrix analysed by fluorescence and thermal imaging. *European Journal of Plant Pathology*, 142(3), 625–632. <https://doi.org/10.1007/s10658-015-0640-9>
- Gutiérrez, S., Diago, M. P., Fernández-Novales, J., & Tardaguila, J. (2018). Vineyard water status assessment using on-the-go thermal imaging and machine learning. *PLoS ONE*, 13(2), 1–19. <https://doi.org/10.1371/journal.pone.0192037>
- Hassan-Esfahani, L., Torres-Rua, A., Jensen, A., & McKee, M. (2015). Assessment of surface soil moisture using high-resolution multi-spectral imagery and artificial neural networks. *Remote Sensing*, 7(3), 2627–2646. <https://doi.org/10.3390/rs70302627>
- Jones, H. G. (1999). Use of infrared thermometry for estimation of stomatal conductance as a possible aid to irrigation scheduling. *Agricultural and Forest Meteorology*, 95(3), 139–149. [https://doi.org/10.1016/S0168-1923\(99\)00030-1](https://doi.org/10.1016/S0168-1923(99)00030-1)
- Khanal, S., Fulton, J., & Shearer, S. (2017). An overview of current and potential applications of thermal remote sensing in precision agriculture. *Computers and Electronics in Agriculture*, 139, 22–32. <https://doi.org/10.1016/j.compag.2017.05.001>

- Kiefer, B., Riemann, M., Büche, C., Kassemeyer, H. H., & Nick, P. (2002). The host guides morphogenesis and stomatal targeting in the grapevine pathogen *Plasmopara viticola*. *Planta*, 215(3), 387–393. <https://doi.org/10.1007/s00425-002-0760-2>
- Lindenthal, M., Steiner, U., Dehne, H. W., & Oerke, E. C. (2005). Effect of downy mildew development on transpiration of cucumber leaves visualized by digital infrared thermography. *Phytopathology*, 95(3), 233–240. <https://doi.org/10.1094/PHYTO-95-0233>
- Linke, M., Beuche, H., Geyer, M., & Hellebrand, H. J. (2000). Possibilities and limits of the use of thermography for the examination of horticultural products. *Agrartechnische Forschung*, 6(2), 110–114. <https://www.cabdirect.org/cabdirect/abstract/20013076270>
- Mahlein, A.-K. (2016). Plant disease detection by imaging sensors- parallels and specific demands for precision agriculture and plant phenotyping. *Plant Disease*, 100(2), 1–11. <https://doi.org/10.1007/s13398-014-0173-7.2>
- Mahlein, A. K., Oerke, E. C., Steiner, U., & Dehne, H. W. (2012). Recent advances in sensing plant diseases for precision crop protection. *European Journal of Plant Pathology*, 133(1), 197–209. <https://doi.org/10.1007/s10658-011-9878-z>
- Martinelli, F., Scalenghe, R., Davino, S., Panno, S., Scuderi, G., Ruisi, P., Villa, P., Stroppiana, D., Boschetti, M., Goulart, L. R., Davis, C. E., & Dandekar, A. M. (2015). Advanced methods of plant disease detection. A review. *Agronomy for Sustainable Development*, 35(1), 1–25. <https://doi.org/10.1007/s13593-014-0246-1>
- O'Shaughnessy, S. A., & Evett, S. R. (2010). Canopy temperature based system effectively schedules and controls center pivot irrigation of cotton. *Agricultural Water Management*, 97(9), 1310–1316. <https://doi.org/10.1016/j.agwat.2010.03.012>
- O'Shaughnessy, Susan A., Evett, S. R., Colaizzi, P. D., & Howell, T. A. (2012). A crop water stress index and time threshold for automatic irrigation scheduling of grain sorghum. *Agricultural Water Management*, 107, 122–132. <https://doi.org/10.1016/j.agwat.2012.01.018>
- Oberti, R., Marchi, M., Tirelli, P., Calcante, A., Iriti, M., & Borghese, A. N. (2014). Automatic detection of powdery mildew on grapevine leaves by image analysis: Optimal view-angle range to increase the sensitivity. *Computers and Electronics in Agriculture*, 104, 1–8. <https://doi.org/10.1016/j.compag.2014.03.001>
- Oerke, E. C., Fröhling, P., & Steiner, U. (2011). Thermographic assessment of scab disease on apple leaves. *Precision Agriculture*, 12(5), 699–715. <https://doi.org/10.1007/s11119-010-9212-3>
- Oerke, E. C., Steiner, U., Dehne, H. W., & Lindenthal, M. (2006). Thermal imaging of cucumber leaves affected by downy mildew and environmental conditions. *Journal of Experimental Botany*, 57(9), 2121–2132. <https://doi.org/10.1093/jxb/erj170>
- Osroosh, Y., Khot, L. R., & Peters, R. T. (2018). Economical thermal-RGB imaging system for monitoring agricultural crops. *Computers and Electronics in Agriculture*, 147(August 2017), 34–43. <https://doi.org/10.1016/j.compag.2018.02.018>

- Osroosh, Y., Troy Peters, R., Campbell, C. S., & Zhang, Q. (2015). Automatic irrigation scheduling of apple trees using theoretical crop water stress index with an innovative dynamic threshold. *Computers and Electronics in Agriculture*. <https://doi.org/10.1016/j.compag.2015.09.006>
- Pedersen, S. M., & Lind, K. M. (2017). *Precision Agriculture : Technology and Economic*. <https://doi.org/https://doi-org.ezproxy.bgu.ac.il/10.1007/978-3-319-68715-5>
- Ray, D. K., Mueller, N. D., West, P. C., & Foley, J. A. (2013). Yield Trends Are Insufficient to Double Global Crop Production by 2050. *PLoS ONE*, 8(6). <https://doi.org/10.1371/journal.pone.0066428>
- Sankaran, S., Mishra, A., Ehsani, R., & Davis, C. (2010). A review of advanced techniques for detecting plant diseases. *Computers and Electronics in Agriculture*, 72(1), 1–13. <https://doi.org/10.1016/j.compag.2010.02.007>
- Savary, S., Ficke, A., Aubertot, J. N., & Hollier, C. (2012). Crop losses due to diseases and their implications for global food production losses and food security. *Food Security*, 4(4), 519–537. <https://doi.org/10.1007/s12571-012-0200-5>
- Seelan, S. K., Laguet, S., Casady, G. M., & Seielstad, G. A. (2003). *Remote sensing applications for precision agriculture : A learning community approach*. 88, 157–169. <https://doi.org/10.1016/j.rse.2003.04.007>
- Shafian, S., & Maas, S. J. (2015). Index of soil moisture using raw Landsat image digital count data in Texas High Plains. *Remote Sensing*, 7(3), 2352–2372. <https://doi.org/10.3390/rs70302352>
- Soliman, A., Heck, R. J., Brenning, A., Brown, R., & Miller, S. (2013). Remote sensing of soil moisture in vineyards using airborne and ground-based thermal inertia data. *Remote Sensing*, 5(8), 3729–3748. <https://doi.org/10.3390/rs5083729>
- Stajanko, D., Lakota, M., & Hočvar, M. (2004). Estimation of number and diameter of apple fruits in an orchard during the growing season by thermal imaging. *Computers and Electronics in Agriculture*, 42(1), 31–42. [https://doi.org/10.1016/S0168-1699\(03\)00086-3](https://doi.org/10.1016/S0168-1699(03)00086-3)
- Stoll, M., Schultz, H. R., Baecker, G., & Berkelmann-Loehnertz, B. (2008). Early pathogen detection under different water status and the assessment of spray application in vineyards through the use of thermal imagery. *Precision Agriculture*, 9(6), 407–417. <https://doi.org/10.1007/s11119-008-9084-y>
- Sullivan, D. G., Shaw, J. N., Mask, P. L., Rickman, D., Guertal, E. A., Luvall, J., & Wersinger, J. M. (2004). Evaluation of multispectral data for rapid assessment of wheat straw residue cover. *Soil Science Society of America Journal*, 68(6), 2007–2013.
- Testi, L., Goldhamer, D. A., Iniesta, F., & Salinas, M. (2008). Crop water stress index is a sensitive water stress indicator in pistachio trees. *Irrigation Science*, 26(5), 395–405. <https://doi.org/10.1007/s00271-008-0104-5>
- Vollmer, M., & Möllmann, K.-P. (2017). Fundamentals of Infrared Thermal Imaging. In *Infrared Thermal Imaging*. <https://doi.org/10.1002/9783527693306.ch1>

- Wen, D. M., Chen, M. X., Zhao, L., Ji, T., Li, M., & Yang, X. T. (2019). Use of thermal imaging and Fourier transform infrared spectroscopy for the pre-symptomatic detection of cucumber downy mildew. *European Journal of Plant Pathology*, 155(2), 405–416. <https://doi.org/10.1007/s10658-019-01775-2>
- Zhou, J., Khot, L. R., Boydston, R. A., Miklas, P. N., & Porter, L. (2018). Low altitude remote sensing technologies for crop stress monitoring: a case study on spatial and temporal monitoring of irrigated pinto bean. *Precision Agriculture*, 19(3), 555–569. <https://doi.org/10.1007/s11119-017-9539-0>

Online References

- Ceramicx, 2011. Ceramix, Retrieved 4 September 2011, from <https://www.ceramicx.com/information/media/white-papers/infrared-heating-for-humans/>
- FLIR, 2019. FLIR from <https://www.flir.com/>
- Shawn Lankton Online, 2007. Shawn Lankton, Active contours Retrieved 15 May 2007, from <http://www.shawnlankton.com/2007/05/active-contours/>

תקציר

תוצרת חקלאית חייבת להכפיל את עצמה עד שנת 2050 בכדי לענות על הביקוש למזון של כלל אוכלוסיית העולם. חקלאות מדייקת מטרתה לתת את הטיפול האופטימלי עבור היחידה הקטנה ביותר בשדה או בחממה. מחלות צמחים מהוות איום גדול על ביטחון המזון העולמי; הינן הגורם המרכזי להפחתה בכמות התוצרת ולהפסדים כלכליים בענף החקלאות. על כן, מעקב וזיהוי מוקדם אחר מחלות בצמחים באופן מהיר ולא הרסני הוא קריטי. כיום, קיים שימוש נרחב בזיהוי מחלות בעזרת ניטור פיזיולוגי של צמחים, אך הוא יקר, כרוך בעבודה רבה ולוקח זמן רב. מידע מוקדם על התפתחות מחלות יכול לשפר את השליטה בהן באמצעות אסטרטגיות ניהול מתאימות.

מטרות מחקר זה הן לזהות כשותית בעלי גפן בשלבי התפתחות מוקדמים באמצעות טכנולוגיית הדמיה תרמית ולקבוע את השעה האופטימלית ביום לצילום התמונות.

בשנים 2019-2020 בוצע ניסוי מבוקר בחממה שבה גודלו צמחי גפן עם עלים נגועים ובריאים. 1403 תמונות תרמיות צולמו 1, 2, 4, 5, 6 ו- 7 ימים לאחר הדבקה מבוקרת. בשנת 2020 בוצע ניסוי מבוקר נוסף על צמחים ותמונות צולמו 6 ימים לאחר הדבקה. 575 תמונות של עלים בריאים ונגועים צולמו בשבעה זמנים שונים בין השעות 7:00 ל- 16:30.

קצוות העלים זוהו באמצעות אלגוריתם המתאר הפעיל. מסכת העלים שנוצרה מהאלגוריתם ומדדים נוספים מהתחנה המטאורולוגית שימשו לחישוב 13 תכונות. באמצעות רגרסיה לוגיסטית בצעדים, זוהו חמשת התכונות משמעותיות ביותר לזיהוי המחלה במודל חיזוי (CWSI2 -I MTD, STD, perc90, Cv).

חמישה מודלים פותחו לסיווג בין עלים נגועים ובריאים: עץ החלטות, רגרסיה לוגיסטית, סיווג בייסיאני נאיבי, מכונת וקטורים תומכים (SVM), והרכבת עצים. הביצועים הוערכו באמצעות אימות צולב (K קבוצות). בוצע ניתוח רגישות באמצעות חלוקות שונות של נתונים על מנת להסביר את התוצאות השונות בין הימים שלאחר ההדבקה. התוצאות הטובות ביותר הושגו על ידי מודל SVM עם דיוק של 81.6%, ציון F1 של 77.5% ו-AUC של 0.874. הזמן האופטימלי לצילום תמונות לזיהוי כשותית היה בין השעות 10:40 ל- 11:30 עם דיוק של 80.7%, ציון F1 של 80.5% ו-AUC של 0.895.

מילות מפתח: חקלאות מדייקת, זיהוי מחלות, סיווג

אוניברסיטת בן-גוריון בנגב
הפקולטה למדעי ההנדסה
המחלקה להנדסה תעשייה וניהול

זיהוי מוקדם של כשותית הגפן בעזרת צילום תרמי

חיבור זה מהווה חלק מהדרישות לקבלת תואר מגיסטר בהנדסה

מאת : בר כהן

מנחים : פרופ' יעל אידן

פרופ' ויקטור אלחנתי

תאריך : 2.10.2021


חתימת המחבר

תאריך : 1.10.2021


אישור המנחים

תאריך : 27.9.2021



תאריך : _____

אישור יו"ר ועדת תואר שני מחלקתית _____

ספטמבר 2021

תשרי תשפ"ב

אוניברסיטת בן-גוריון בנגב
הפקולטה למדעי ההנדסה
המחלקה להנדסה תעשייה וניהול

זיהוי מוקדם של כשותית הגפן בעזרת צילום תרמי

חיבור זה מהווה חלק מהדרישות לקבלת תואר מגיסטר בהנדסה

מאת : בר כהן

ספטמבר 2021

תשרי תשפ"ב