

2018

## Detection of multi-tomato leaf diseases (*late blight, target and bacterial spots*) in different stages by using a spectral-based sensor

Jinzhu Lu  
Xihua University, [lujingzhu1103@163.com](mailto:lujingzhu1103@163.com)


Reza Ehsani  
University of California, Merced

Yeyin Shi  
University of Nebraska-Lincoln, [yshi18@unl.edu](mailto:yshi18@unl.edu)

Ana Isabel de Castro  
Institute for Sustainable Agriculture

Shuang Wang  
Xihua University

Follow this and additional works at: <http://digitalcommons.unl.edu/biosysengfacpub>

 Part of the [Bioresource and Agricultural Engineering Commons](#), [Environmental Engineering Commons](#), and the [Other Civil and Environmental Engineering Commons](#)


---

Lu, Jinzhu; Ehsani, Reza; Shi, Yeyin; de Castro, Ana Isabel; and Wang, Shuang, "Detection of multi-tomato leaf diseases (*late blight, target and bacterial spots*) in different stages by using a spectral-based sensor" (2018). *Biological Systems Engineering: Papers and Publications*. 522.

<http://digitalcommons.unl.edu/biosysengfacpub/522>

This Article is brought to you for free and open access by the Biological Systems Engineering at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Biological Systems Engineering: Papers and Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

# SCIENTIFIC REPORTS



OPEN

## Detection of multi-tomato leaf diseases (*late blight*, *target* and *bacterial spots*) in different stages by using a spectral-based sensor

Jinzhu Lu<sup>1</sup>, Reza Ehsani<sup>2</sup>, Yeyin Shi<sup>3</sup>, Ana Isabel de Castro<sup>4</sup> & Shuang Wang<sup>1</sup>

Several diseases have threatened tomato production in Florida, resulting in large losses, especially in fresh markets. In this study, a high-resolution portable spectral sensor was used to investigate the feasibility of detecting multi-diseased tomato leaves in different stages, including early or asymptomatic stages. One healthy leaf and three diseased tomato leaves (*late blight*, *target* and *bacterial spots*) were defined into four stages (healthy, asymptomatic, early stage and late stage) and collected from a field. Fifty-seven spectral vegetation indices (SVIs) were calculated in accordance with methods published in previous studies and established in this study. Principal component analysis was conducted to evaluate SVIs. Results revealed six principal components (PCs) whose eigenvalues were greater than 1. SVIs with weight coefficients ranking from 1 to 30 in each selected PC were applied to a *K*-nearest neighbour for classification. Amongst the examined leaves, the healthy ones had the highest accuracy (100%) and the lowest error rate (0) because of their uniform tissues. Late stage leaves could be distinguished more easily than the two other disease categories caused by similar symptoms on the multi-diseased leaves. Further work may incorporate the proposed technique into an image system that can be operated to monitor multi-diseased tomato plants in fields.

Fresh-market tomatoes are produced in every state in the US, where 20 states support commercial-scale production. According to data from the US Department of Agriculture, Florida is the leading state in fresh-market tomato production. However, a series of diseases has threatened tomato production in Florida, resulting in large losses in fresh and processed tomato production. Tomato diseases are caused by several factors, including fungal, bacterial and viral infections<sup>1–4</sup>. Most of the foliar diseases, such as *late blight*, *target* and *bacterial spots*, are favoured by warm temperatures or prolonged periods of wetness, which are typical in Florida. Conventional scouting for foliar diseases relies primarily on the visual inspection of leaf colour patterns and crown structures. Laboratory test approaches, such as polymerase chain reaction, enzyme-linked immunosorbent assay and loop-mediated isothermal amplification, are highly specific and sensitive to identify diseases on plant tissue samples. However, these approaches involve destructive methods, entail time- and labour-consuming processes and require specialised skills<sup>5–8</sup>.

Advancements in agricultural technology have offered opportunities for the non-destructive detection of plant diseases through spectroscopy<sup>9</sup>. A visible near-infrared spectrometer (400–1000 nm) is used in a reflectance mode to distinguish sprouted and intact wheat kernels under laboratory conditions<sup>10</sup>. Reflectance at 728 and 878 nm is utilised to classify sprouted and intact kernels, and the wavelength region above 720 nm is set to categorise sprouted kernels according to different levels of severity. The correct recognition rates of intact, sprouted and severely sprouted kernels are 100%, approximately 94% and 98%, respectively. The ability of reflectance spectroscopy in three regions, namely, ultraviolet, visible and near-infrared, has been evaluated indoors to determine the disease severity of tomato leaves infected with *Xanthomonas perforans*<sup>11</sup>. Partial least squares (PLS) regression, stepwise multiple logistic regression (SMLR) and their combinations have been applied to derive four predictive

<sup>1</sup>School of Mechanical Engineering, Xihua University, 999 Jinzhou Road, Chengdu, Sichuan, 610000, China.

<sup>2</sup>Mechanical Engineering Department, University of California-Merced, 5200 N. Lake Road, Merced, CA, 95343, U.S.A. <sup>3</sup>Department of Biological Systems Engineering, University of Nebraska-Lincoln, 3605 Fair Street, Lincoln, NE, 68583, United States. <sup>4</sup>Department of Crop Protection, Institute for Sustainable Agriculture (IAS-CSIC), Cordoba, Spain. Correspondence and requests for materials should be addressed to J.L. (email: [lujingzhu1103@163.com](mailto:lujingzhu1103@163.com))

models. Amongst these four models, the model established by SMLR is the most efficient in predicting disease severity with a root mean square difference of 4.9% and a coefficient of determination of 0.82. Ground-level reflectance spectra have also been obtained for the in-field detection of plant nitrogen<sup>12–14</sup>. Spectral signatures of spectral data have been analysed in leaves to differentiate sugar beet diseases<sup>15</sup>. Spectral reflectance has been measured in fields by using a handheld spectroradiometer in the range of 400–1050 nm, and the highest correlation coefficients ( $r = 0.85$ ) have been detected in the visible region. This study has also provided a basis for further development of classification methods of sugar beet diseases in different developmental stages.

Spectral vegetation indices (SVIs) from ground-level reflectance spectral data can be used to estimate crop yield<sup>16</sup>, detect variations in leaf area index<sup>17</sup> and characterise agricultural crop biophysical variables<sup>18</sup>. SVIs can also be used to detect or differentiate various plant diseases<sup>19–21</sup>. Different diseases are often associated with specific physiological and visual changes in their host plants. Appropriate SVIs offer a great advantage in the dimension reduction of spectral data<sup>22</sup>. Spectral studies on detecting foliar diseases have been performed using destructive chemical methods<sup>23–25</sup>. Some studies have reported the use of non-destructive methods to detect foliar diseases on certain varieties. Few studies have focused on multi-diseased leaves with similar symptoms regardless of plant variety. Spectral data can be obtained using spectral sensors and hyperspectral imaging sensors. In agriculture, most spectral sensors are utilised to detect the internal qualities of fruits, such as sugar degree. Hyperspectral images are obtained to detect plant diseases, especially foliar diseases. Considering that diseases may cause symptoms on leaves, we may examine images to obtain a visual leaf image, especially when multiple scars are present on one leaf. Hyperspectral imaging technology requires large and time-consuming computation. Thus, this imaging is usually performed in laboratories for research involving detection approaches. In our previous work, we improved the possibility of using spectroscopy technology in plant disease detection. In the present study, we aimed to perform an in-depth research on the detection of multi-diseased leaves and promote spectroscopy technology on the detection of leaf diseases.

This research generally aimed to investigate the feasibility of a simple method to detect multi-diseased tomato leaves in different stages, including an early or asymptomatic stage, by using a spectral-based sensor. This study was specifically designed 1) to calculate and create SVIs from reflectance spectra that can describe healthy and diseased leaves in different stages; 2) to select SVIs consisting of multi-wavelengths that can mostly describe changes in leaves and 3) to establish and evaluate classification models on differentiating healthy, asymptomatic, early stage and late stage diseased plants by using the selected SVIs.

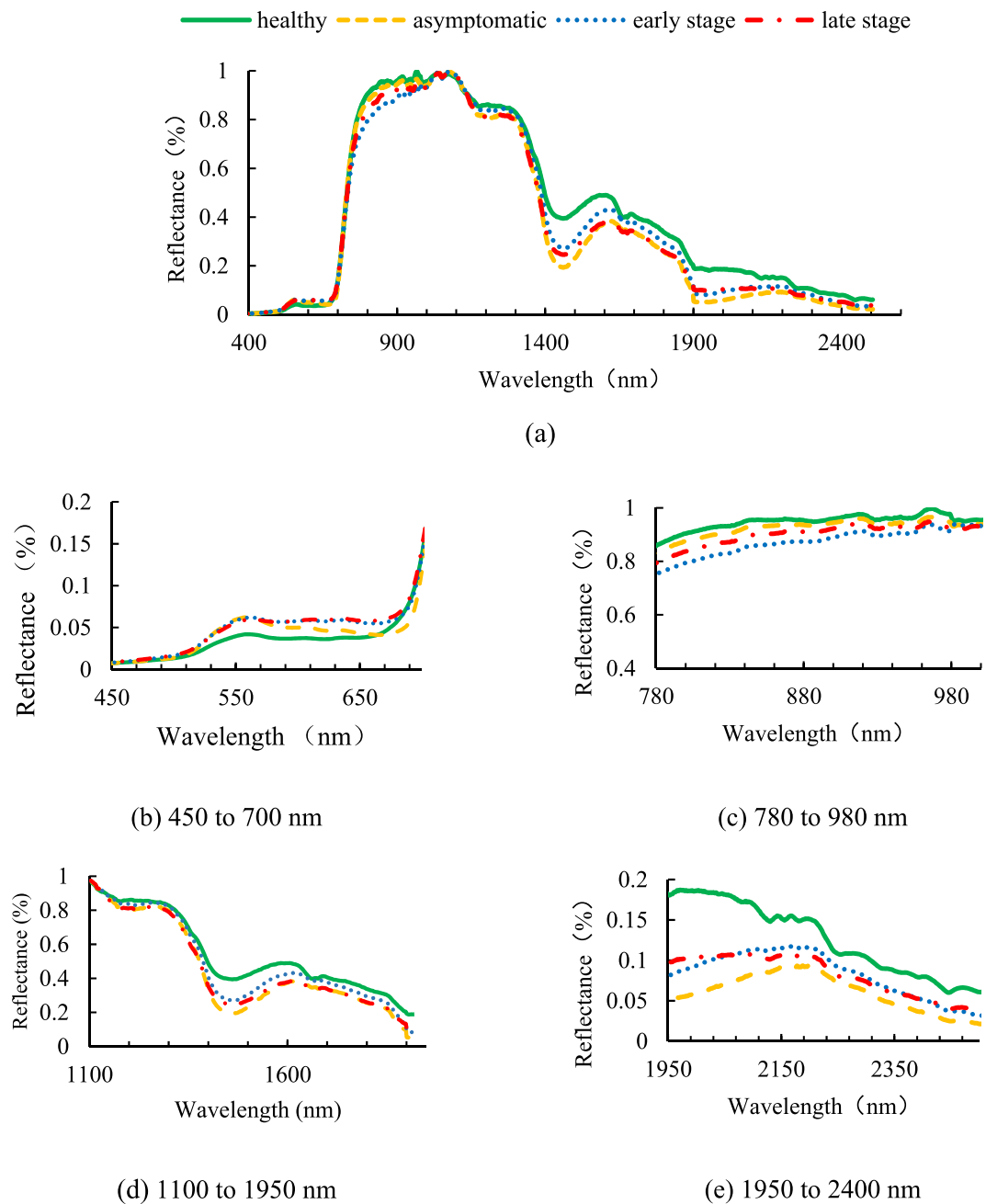
## Results

**Spectra of leaves.** Figure 1 shows the averaged reflectance spectra of all of the data after a baseline correction was performed [Equation (1)]. These four spectra significantly differ in several bands in the whole wavelength (Fig. 1). To observe the spectra in detail, we zoomed in and divided four reflectance spectral bands into one visible band and three near-infrared bands [Fig. 1(b) to (e)].

Figure 1(b) illustrates the zoomed reflectance spectra at 450–700 nm. Compared with the spectrum of the healthy leaf, the spectra of three other leaf diseases increased in reflectance from 500 nm. In the range of 500–570 nm, these three curves were intersected. Beyond 570 nm, two symptomatic diseased leaves ('early stage' and 'late stage') remained intersected, and both were higher than the two other categories. The average reflectance of the asymptomatic leaves decreased from 570 nm and then kept lower than that of the two symptomatic leaves. The reflectance of the asymptomatic leaves was retained at levels higher than that of the healthy ones from 660 nm and beyond. All of the spectra increased suddenly at around 690 nm. Figure 1(c) shows the zoomed reflectance spectra at 780–1000 nm. The reflectance spectra of the four categories represented regular changes. The reflectance spectra of healthy, asymptomatic, early stage and late stage were in a decreasing trend. All of the spectra exhibited a flattened slope, whose reflectance was not smaller than 75% and not greater than 100%. Figure 1(d) illustrates the zoomed in reflectance spectra at 1100–1950 nm. All of the reflectance spectra declined from 1100 nm to around 1400 nm in one bottom section in the region. The change in the four spectra was gradually irregular. Figure 1(e) shows the zoomed in reflectance spectra in 1950–2400 nm. The reflectance spectra of the healthy leaves were significantly higher than those of the three diseased leaves. The spectral curves of the asymptomatic and early stage leaves showed a similar trend, which peaked at around 2200 nm, and declined slowly until 2500 nm was reached. The spectra of the late stage leaves had two intersection points with those of the early stage leaves at 2050 and 2390 nm.

**Spectral vegetation indices by PCA.** Fifty SVIs from previous studies and seven SVIs obtained in this study were subjected to principal component analysis (PCA). Their eigenvalues were used as assessment criteria. An eigenvalue of less than 1 for one principal component (PC) of an SVI indicates that the effect of such a PC is less important than single variable. Therefore, only PCs with eigenvalues greater than 1 were chosen for the next step in this study. In each chosen PC, the SVIs were arranged according to their weight coefficient starting from the highest to the lowest. A total of 147 samples were included in one category in this test. Considering the principle 'sample number is 5–10 times greater than the optimal feature number', we used the minimum and maximum numbers of features as two cut points, whose components were scored in a column with a list of the first 15 and 30. Table 1 shows the PCs and the spectral indices combinations. Furthermore, 6 eigenvalue roots of the PCs were greater than 1. In each of the 6 PCs, 15 and 30 SVIs were found. The SVIs with the highest frequency were  $PSSR_{\alpha}$ ,  $NDWI-Hyp$ ,  $mSRI_3$  and  $BRI_2$ . The SVIs with a frequency of less than 1 were  $SIP1$ ,  $ND_5$ ,  $NDVI$ ,  $ND_1$ ,  $SR_1$  and  $RDVI$ . Considering the low frequency, we found that the SVIs were slightly correlated with classification, and only SVIs ranking from 1 to 30 were selected as input in the classifier for further data analysis.

**Classification.** For each sample, two SVI combinations were used for classification. One combination involved SVIs ranking from 1 to 15, and the other combination comprised SVIs ranking from 1 to 30. The



**Figure 1.** Average reflectance spectra of healthy, asymptomatic, early stage and late stage leaves: (a) spectral ranges of 400–2500 nm, (b) zoomed in spectral ranges of 450–700 nm, (c) zoomed in spectral ranges of 780–980 nm, (d) zoomed in spectral ranges of 1100–1950 nm and (e) zoomed in spectral ranges of 1950–2400 nm.

classification results of the two datasets were compared to evaluate the effects of the number of SVIs. Both datasets were divided into training and testing groups based on the ratio of 70%:30%. For instance, the error rate of ‘C’ was C samples misclassified into other categories divided by all C samples. The accuracy of ‘C’ is correctly classified as C samples in all of the C samples. The training set was treated as a known group because KNN used lazy learning, and the classification results consisted of accuracy and error rate based on the testing group. Table 2 summarises the error rate of the healthy, asymptomatic, early stage and late stage leaves. Based on the error rate of the testing group, the highest error rate of the four categories with different PCs resulted in 30 SVIs. In the six PCs, the lowest error rates of healthy, asymptomatic, early stage and late stage leaves did not have significant differences, which were 0, 4.8, 5.6 and 0, respectively. However, classification difficulties observed when 30 SVIs of healthy, asymptomatic, early stage and late stage leaves were used, and the error rates were 26.9% in PC1, 28.6% in PC1 or PC3, 23.3% in PC4 and 25% in PC1. The asymptomatic leaves could be classified as the worst amongst the three other leaves, while the three other categories were categorised correctly at least in one PC combination.

| PCs | Eigenvalue | SVIs rank 1–15                                                                                                                                                                                                                             | SVIs rank 16–30                                                                                                                                                                                                                                                   |
|-----|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| PC1 | 25.521     | ND <sub>1</sub> , CTR <sub>3</sub> , ND <sub>3</sub> , SR <sub>3</sub> , SIPI, SR <sub>6</sub> , NDVI, MSR <sub>2</sub> , SR <sub>5</sub> , SR <sub>4</sub> , ND <sub>2</sub> , ND <sub>3</sub> , ND <sub>4</sub> , mND, mSRI <sub>3</sub> | LIC, BRI <sub>1</sub> , DSWI <sub>5</sub> , VOG, SR <sub>7</sub> , PSSR <sub>3</sub> , NDWI, WI, RGRI, BRI <sub>2</sub> , PSSR <sub>3</sub> , BI, CTR <sub>1</sub>                                                                                                |
| PC2 | 13.880     | BGI <sub>1</sub> , SR <sub>1</sub> , BGI <sub>3</sub> , SR <sub>2</sub> , CI <sub>1</sub> , DSWI <sub>2</sub> , RDVI, MTVI, AVE, TVI, DSWI <sub>4</sub> , DSWI <sub>3</sub> , MCAI, CI <sub>2</sub> , GREEN I                              | PRI <sub>2</sub> , PSSR <sub>3</sub> , REP, PRI <sub>1</sub> , PRI <sub>3</sub> , TCARI, RGRI, MCARI, BI, NDWI-Hyp, PSSR <sub>3</sub> , VOG, SR <sub>7</sub> , BRI <sub>2</sub> , ND <sub>2</sub>                                                                 |
| PC3 | 6.398      | ARI, MCARI, TCARI, MSI, TVI, DSWI <sub>2</sub> , MCAI, AVE, MTVI, DSWI <sub>1</sub> , REP, RDVI, NDWI, WI, NDWI-Hyp                                                                                                                        | DSWI <sub>5</sub> , SR <sub>1</sub> , PSSR <sub>3</sub> , SR <sub>3</sub> , BRI <sub>2</sub> , PSSR <sub>3</sub> , CTR <sub>1</sub> , RI, RGRI, mSRI <sub>1</sub> , mSRI <sub>3</sub> , BGI <sub>1</sub> , BRI <sub>1</sub> , SR <sub>5</sub> , mSRI <sub>2</sub> |
| PC4 | 5.194      | RI, CI <sub>2</sub> , CTR <sub>1</sub> , PRI <sub>1</sub> , SR <sub>7</sub> , VOG, RE, CI <sub>1</sub> , BRI <sub>2</sub> , GREEN I, NDWI-Hyp, PSSR <sub>3</sub> , BGI <sub>2</sub> , MSR <sub>3</sub> , MSR <sub>1</sub>                  | mND, SR <sub>4</sub> , ND <sub>3</sub> , ND <sub>4</sub> , REP, ARI, MCARI, DSWI <sub>3</sub> , DSWI <sub>4</sub> , SR <sub>5</sub> , TCARI, RDVI, LIC, MTVI, mSRI <sub>2</sub>                                                                                   |
| PC5 | 2.057      | NDWI-Hyp, DSWI <sub>1</sub> , CTR <sub>1</sub> , BRI <sub>1</sub> , WI, BRI <sub>2</sub> , MSI, NDWI, DSWI <sub>5</sub> , PRI <sub>3</sub> , PRI <sub>2</sub> , LIC, RGRI, MSR <sub>3</sub> , MSR <sub>1</sub>                             | DSWI <sub>3</sub> , DSWI <sub>4</sub> , ARI, SR <sub>1</sub> , SR <sub>2</sub> , GREEN I, MCAI, PSSR <sub>3</sub> , TVI, SR <sub>3</sub> , CTR <sub>2</sub> , MTVI, AVE, BI, CI <sub>1</sub>                                                                      |
| PC6 | 1.080      | BI, PRI <sub>2</sub> , LIC, NDWI, PRI <sub>3</sub> , BGI <sub>1</sub> , MCAI, RI, DSWI <sub>1</sub> , RE, VOG, SR <sub>7</sub> , BGI <sub>2</sub> , WI, SR <sub>4</sub>                                                                    | ND <sub>4</sub> , ND <sub>3</sub> , BRI <sub>1</sub> , CI <sub>1</sub> , mND, DSWI <sub>5</sub> , MSR <sub>1</sub> , MSR <sub>3</sub> , TVI, MTVI, AVE, MSI, CTR <sub>2</sub> , NDWI-Hyp, RDVI                                                                    |

**Table 1.** PCs with different eigenvalues and their component SVIs ranking from 1 to 30.

| PCs | Healthy |        | Asymptomatic |        | Early stage |        | Late stage |        |
|-----|---------|--------|--------------|--------|-------------|--------|------------|--------|
|     | 15SVIs  | 30SVIs | 15SVIs       | 30SVIs | 15SVIs      | 30SVIs | 15SVIs     | 30SVIs |
| PC1 | 11.8    | 26.9   | 9.5          | 28.6   | 19.5        | 7.3    | 1          | 25     |
| PC2 | 0       | 0      | 19           | 16.1   | 7.7         | 12.5   | 7.9        | 4.7    |
| PC3 | 8.3     | 0      | 9.5          | 28.6   | 5.6         | 13.9   | 9.5        | 0      |
| PC4 | 25      | 0      | 17.8         | 13     | 0           | 23.3   | 15         | 0      |
| PC5 | 12.5    | 0      | 21.7         | 4.8    | 6.8         | 10.2   | 2.9        | 2.3    |
| PC6 | 0       | 0      | 9.5          | 25     | 11.1        | 7.8    | 5.6        | 8.3    |

**Table 2.** Error rates of classifying healthy, asymptomatic, early stage and late stage leaves with six PCs and two SVI compositions.

| PCs | Healthy |        | Asymptomatic |        | Early stage |        | Late stage |        |
|-----|---------|--------|--------------|--------|-------------|--------|------------|--------|
|     | 15SVIs  | 30SVIs | 15SVIs       | 30SVIs | 15SVIs      | 30SVIs | 15SVIs     | 30SVIs |
| PC1 | 100     | 90.5   | 95           | 65.2   | 84.6        | 75     | 77.1       | 90.9   |
| PC2 | 95.7    | 100    | 89.5         | 92.9   | 92.3        | 85.4   | 89.7       | 91.1   |
| PC3 | 100     | 100    | 86.4         | 87.0   | 89.5        | 81.6   | 92.7       | 91.4   |
| PC4 | 85.7    | 100    | 100          | 76.9   | 73.5        | 94.3   | 100        | 85.3   |
| PC5 | 100     | 100    | 90.0         | 83.3   | 85.4        | 97.8   | 91.7       | 95.5   |
| PC6 | 100     | 100    | 86.4         | 85.7   | 93.0        | 90.4   | 91.9       | 86.8   |

**Table 3.** Classification results of healthy, asymptomatic, early stage and late stage leaves with six PCs and two SVIs compositions.

Table 3 summarises the accuracy of healthy, asymptomatic, early stage and late stage leaves. Each of the six PCs contained two SVI combinations. A total of 12 results were included in one category. Based on the accuracies of the testing group, the highest accuracies of the healthy, asymptomatic and late stage leaves were 100%. By comparison, the highest accuracies of different PC or SVIs of early stage leaves ranged from 73.5% to 97.8%. In different PCs or two SVI combination groups, the lowest accuracies of the healthy, asymptomatic, early stage and late stage leaves were 85.7%, 65.2%, 73.5% and 77.1%, respectively. The difficulties in the classification of the asymptomatic leaves occurred in PC1 when 30 SVIs were used, and the accuracy was 65.2%. In the three other categories, the accuracies were higher than 70%.

## Discussion

**Spectral analysis.** In the visible range, the diseased leaves, especially those with symptoms, had spots or scars. The symptoms of these three diseases are light or dark brown spots on the leaf. The colour induced less absorption in the diseased leaves in the reflectance mostly in the green range (495–570 nm) and in the red edge range (around 690 nm). As an indicator of plant chlorophyll content<sup>26,27</sup>, the red edge (680–750 nm) differed in the spectrum slope. The effect of disease on leaf wetness, which indicates the relationship of infection with water content, has been studied<sup>28,29</sup>. Laboratory studies on leaf water content have been conducted in the near infrared region of the leaf spectra (780–2500 nm)<sup>30–32</sup>. The average spectra of the asymptomatic and healthy leaves were compared, and the results demonstrated that the two other symptomatic leaves exhibited a strong spectral change in the entire spectral region.

In the spectral range of 1000 nm and above, the reflectance of the four categories of leaves did not significantly change, and this condition may be observed in multiple diseases. Some studies have reported that plant cellular contents, such as enzymes and proteins, which represent the changes in the reflectance spectra in the range of

1000–2500 nm, are affected by specific pathogens<sup>33</sup>. Each pathogen causes specific changes in the spectra that show regular pathogenic patterns. In our experiments, the diseased leaves were collected from a field and inoculated with pathogens in the natural environment. A certain pathogen in each leaf was difficult to control, and the disease consequently spread in one area. Enzymes elicit various effects on different diseases. Moreover, we classified the leaves according to the disease stage instead of the varieties of the diseases. As such, changes in the multi-diseased leaves were irregular in the spectral range exceeding 1000 nm.

**SVI calculation.** The majority of SVIs were sourced from previous studies, whereas seven SVIs were established in this study based on the analysis of the spectral reflectance. After PCA was conducted, each of the 30 SVIs of the 6 PC combinations was selected. Four SVIs with the highest frequency (PSSR<sub>c</sub>, NDWI-Hyp, mSRI<sub>3</sub>, BRI<sub>2</sub>) and six SVIs (SIPI, ND<sub>5</sub>, NDVI, ND<sub>1</sub>, SR<sub>1</sub> and RDVI) were not observed in different PC combinations. These high-frequency SVIs were associated with the wavelengths in the green region (445 and 450 nm), red and far-red region (690, 707 and 750 nm) and near-infrared region (800, 1070 and 1200 nm). However, the SVIs that were not selected for PCA had the same wavelengths at 800 nm (SIPI, NDVI, ND<sub>1</sub>, SR<sub>1</sub> and RDVI) and 750 nm (ND<sub>5</sub>). In these two wavelengths, significant differences in the spectra were found between the healthy and diseased leaves and between the leaves in two disease stages. However, the spectra in this study were all averaged in at least 74 spectra. Several spectra were detected above or below the average. Therefore, one disease or even multi-diseased leaves could be inadequately described on single wavelengths. This study also explained why two SVIs (ND<sub>5</sub> and mSRI<sub>3</sub>) created by some wavelengths in this study differed from the selected SVIs. High-frequency SVIs, such as PSSR<sub>c</sub>, mSRI<sub>3</sub> and BRI<sub>2</sub>, were related to pigments, such as carotenoids and chlorophyll. Visual diseased spots on the leaves were caused by pigment losses. However, asymptomatic leaves varied from the expected finding except the disrupted plant cells<sup>34</sup>, and several diseases affect leaves even in the asymptomatic stage. Low-frequency SVIs, such as NDVI, SIPI and RDVI, and spectral ratios are highly related to plant vigour or pigments in remote sensing studies<sup>35,36</sup>, especially on nitrogen detection or soil status by SVIs<sup>37</sup>. In terms of foliar disease stages, these SVIs were poorly correlated with classification in this study.

**Classification.** When the plants were affected by diseases in this study, especially in the late stage, visual spots or scars were found on the leaves. Amongst the diseased leaves, the late stage leaves showed the highest classification accuracy and were the least affected by the number of two SVIs, whose accuracy was 14.7%. In experiment sampling, regions of interests (ROIs) were selected mainly around the lesions. The heavier the tissue was, the greater the differences between healthy tissues and diseased spots would be. Therefore, the late stage leaves could be distinguished more easily than the two other disease categories. The accuracies of the early stage and asymptomatic leaves were lower than those of the leaves in other stages. The accuracy of the asymptomatic leaves was higher than that of the early stage leaves, whereas the error rate of the asymptomatic leaves was lower than that of the other leaves. This misjudgement might be attributed to the confusion between asymptomatic and healthy leaves, that is, when the plants were initially affected by the disease, the viruses or bacteria existed in tissues or cells. Pathogens usually undergo a long incubation period before symptoms can be observed. Amongst the four categories in this study, the healthy leaves had the highest accuracy (100%) and the lowest error rate (0) mainly because of sampling. The spectroscopy sensor had a 4° field of view (FOV), and the distance between the sensor and the leaves was about 50 cm. Therefore, the ROIs could not be concentrated at a certain point. When the healthy leaves were sampled, the tissues were all healthy regardless of ROIs. By comparison, the regions of the tissues of the healthy and diseased areas on the diseased leaves overlapped. Consequently, the spectral data of the diseased leaves did not completely represent the diseased categories. Nevertheless, the highest accuracy of classifying the healthy, asymptomatic and late stage leaves was 100% when different PCs and SVIs were used, thereby confirming the feasibility of using the spectroscopy-based sensor with SVIs for the classification of tomato multi-disease leaves with similar symptoms even in the asymptomatic stage. Further studies should incorporate the proposed technique into an image system that can be operated to monitor multi-diseased tomato plants in fields.

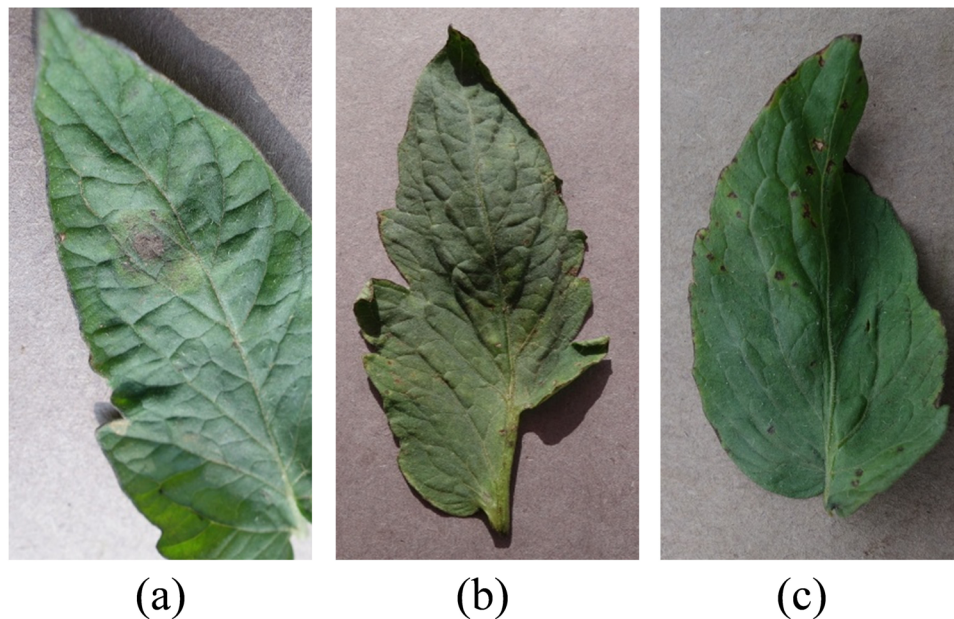
## Methods

**Spectral Sensor.** A portable high-resolution SVC HR-1024 spectroradiometer (Spectra Vista Corporation, Poughkeepsie, NY, USA) was used to collect reflectance data in the range of 350–2500 nm, with spectral resolutions of less than or equal to 3.5, 9.5 and 6.5 nm at the wavelength ranges of 350–1000, 1000–1850 and 1850–2500 nm, respectively. The spectral data were collected using a 4° FOV lens at a minimum integration time of 4 ms. To acquire the relative reflectance spectra of the sample, we conducted a correction by using Equation (1). The raw spectra ( $R_{\text{sample}}$ ) were calibrated in terms of white ( $R_{\text{reference}}$ ) and dark correction ( $R_{\text{dark}}$ ). The white correction was obtained with a white reference panel (Spectralon Reflectance Target, CSTM-SRT-99-100 Spectra Vista Corporation, Poughkeepsie, NY, USA), and the dark correction was carried out by covering the lens with a light-proof cap.

$$R = \frac{R_{\text{sample}} - R_{\text{dark}}}{R_{\text{reference}} - R_{\text{dark}}} \quad (1)$$

**Diseases.** *Late blight*, *target* and *bacterial spot* are three widely diseases that spread in tomatoes produced in Florida. In Fig. 2, all of these diseases manifested similar symptoms: irregular yellow necrotic areas, irregularly shaped water-soaked or brown spots and lesions on tomato leaves. The infected plants manifest necrotic crowns and die<sup>38</sup>. Infection caused by one pathogen may cause plants to become susceptible to infection or colonization





**Figure 2.** Three diseased tomato leaves with disease symptoms: (a) *Late blight*; (b) *Target*; (c) *Bacterial spot*. Original source: Plant Village<sup>43</sup>.

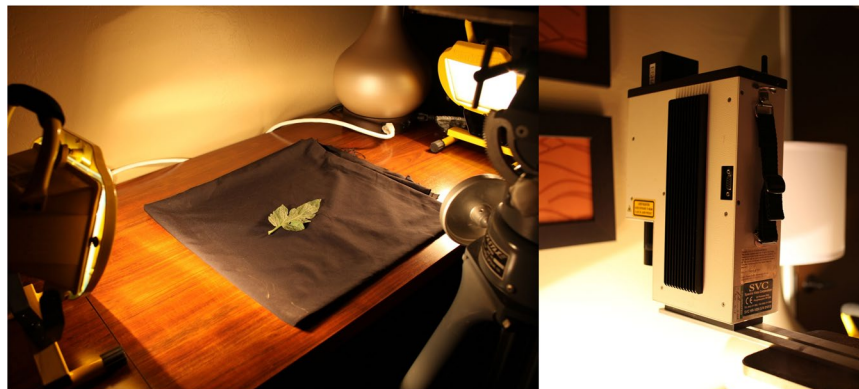
by other pathogens, leading to multi-diseased leaves. With these adverse conditions, tomato diseases should be detected in early stages in fresh markets.

**Sampling.** Experiment was conducted in a commercial farm in Naples, Florida. All of the leaves were collected by professional experts, who defined the diseased stages: ‘asymptomatic’ (leaves from the diseased plant without foliar symptoms that can be observed by human eyes), ‘early stage’ (symptomatic leaves from diseased plants that can be cured by fungicides) and ‘late stage’ (symptomatic leaves from incurable diseased plants). With warm temperatures or periods of wetness in Florida, diseases spread fast and induce cross infections. Multi-diseased leaves are commonly seen in actual fields. Other studies on tomato diseases usually involve artificial inoculation method or strict detachment of leaves to ensure that only one disease affects one leaf. In our experiment, multi-diseased leaves were selected in an uncontrolled cultivation environment, which usually spread three diseases, namely, *late blight*, *bacterial spot* and *target*. Multi-diseased leaves from plants in each disease stage have one or two characteristic symptoms of a particular disease. Therefore, we classified the leaves based on leaf stage instead of disease varieties. *Late blight* was observed in the asymptomatic stage, *bacterial spot* and *late blight* were detected in early stage, and *bacterial spot* and *target* were found in late stage.

In each stage, three or five leaves were collected from one plant in each disease stage. Each leaf sample was placed in an isolated plastic bag immediately stored in an icebox. Measurements were performed within 3 h after the leaves were collected. The proposed detection system was composed of the sensor and two 500 W halogen portable lights from two symmetric sides to provide a uniform light condition (Fig. 3). At the start of each measurement, other lights were turned off to ensure that no any interference light was present during the experiment. The leaves were placed in a fixed plate in a black background. The sensor was installed 45–50 cm high from the plate. Three measurements were conducted in one leaf containing three different spots on one leaf to obtain an average spectrum in the dataset. After sampling was completed, the spectra were checked manually. Any spectrum with interfering signals was removed before it was used as an input of the dataset. The spectral data of the corresponding sample were collected again. The dataset consisted of 74, 77, 148 and 146 leaf spectra for healthy, asymptomatic, and symptomatic categories, respectively.

**Spectral vegetation indices.** SVI is used to enhance the characteristics of information, such as the object’s properties, based on two or more wavelengths or range of wavelength reflectance combination. Seven kinds of vegetation indices, namely, broadband of green, narrow strip of green, light canopy of utilization, nitrogen, drought or carbon decay, pigment and canopy water content, were achieved. We listed 50 normally used SVIs and 7 SVIs created in this study (Table 4). These SVIs in the wavelengths of the visible and near-infrared regions were calculated from the raw spectra after baseline correction and outlier removal were performed.

**Principal Component Analysis (PCA).** PCA is a commonly used multivariate statistical method that simplifies datasets. In orthogonal transformation, the observations of possibly correlated variables were converted into a set of values of linearly uncorrelated variables called principal components (PC). In geometry, this transformation is equivalent to coordinating selection or translation in a multidimensional vector space. In algebra, this transformation corresponds to solving a feature value of a covariance matrix to obtain new variables (axis). According to the size order of variance (feature value), these variables become the first main components (PC1),



**Figure 3.** Detection system in our laboratory.

the second main components (PC2), the third main components (PC3), and so on, indicating that PC1 has the largest possible variance. The main idea of PCA is data dimension reduction. The number of new variables is less than or equal to the number of original variables. In PCA, a weight coefficient of one new variable represents the importance of the new variable.

PCA is commonly used to select sensitive wavelengths in spectral analysis. However, according to its principle, one component (PC) consists of more than one wavelength. Using one wavelength may lack persuasiveness for further classification. In this study, combining SVIs and PCA could solve the problem. The flowchart of data processing is shown in Fig. 4. We subjected the calculated 57 SVIs to PCA and used the eigenvalue as the assessment criteria. Only PCs with eigenvalues greater than 1 were then chosen for the next step. In each selected PC, the SVIs were arranged according to their weight coefficient starting from the highest to the lowest. Each PC consisted of 57 variables (SVIs) with an individual weight coefficient that can represent the importance of one variable. However, using one variable is similar to utilising one wavelength. As such, it may lack persuasiveness for further classification. Therefore, numerous SVIs, which consisted of multi-wavelengths, were selected in one PC in this work. Furthermore, for each sample, two SVI combination groups were applied to compare the classification. According to the principle, 'sample number is 5–10 times greater than the optimal feature number', the minimum and maximum numbers of features were treated as two cut points, whose components were scored in a column that was listed for each PC. Thus, the effects of the number of SVIs can be evaluated in terms of their classification results.

**K-Nearest Neighbour (KNN).** Classification analyses were conducted in PASW Statistics 18 (The SPSS Inc., Quarry Bay, Hong Kong). The *K*-nearest neighbour (KNN) used in this study is a simple classifier that works well on basic recognition problems in machine learning techniques<sup>39,40</sup>. Classification is achieved by identifying the nearest neighbours to a query example and then using such neighbours to determine the class of the query. This method is easily implemented and can obtain good results if the neighbours (*k* value) are chosen carefully.

In this work, we obtained a small dataset of 74, 77, 148 and 146 healthy, asymptomatic, early and late stage categories of diseased tomato leaves, respectively. According to our previous work<sup>41,42</sup>, KNN is an optimal classifier for the detection of diseases in indoor plants with a sample size of less than 200. Considering the continuity and the objectives of our work, KNN was used for further classification.

In the first step, the sample distance between the unspecified group and the known group is calculated. Then, the *k* neighbours of the unspecified samples at a certain distance are located. If these *k* neighbours belong to one group, the sample is assigned to the category. If more than one group is found in the *k* neighbours, the sample is included in the major category. In this paper, the sample distance was calculated by Mahalanobis distance method because the samples were numerical [Equation (2)]:

$$d(X_i, X_j) = \|X_i - X_j\| = \left( \sum_k^m W_k |X_{ik} - X_{jk}|^2 \right)^{\frac{1}{2}} \quad (2)$$

where  $X_i$  and  $X_j$  are two samples in the dataset;  $n$  is the number of sample properties;  $W$  is the category in which  $X_i$  or  $X_j$  is assigned.

When *k* has a different value, the results change. A small *k* value may cause overfitting amongst the data in the model, whereas a large *k* value requires a long computation time. The *k* values selected in this study were 3, 4 and 5. The best *k* value was determined through cross-validation.

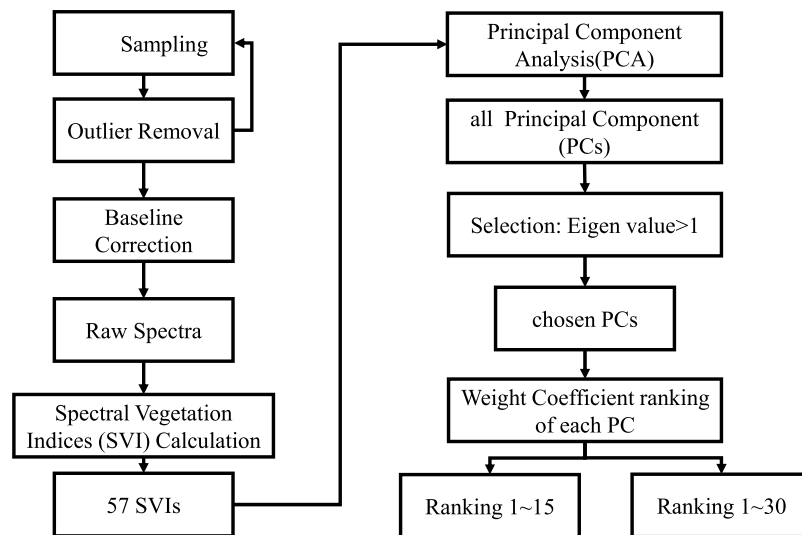
**Classification.** The accuracy and error rate are determined by equations (3) and (4).

$$\text{accuracy} = \frac{\text{correctly classified leaves samples in } C}{\text{all leaves samples in } C} \quad (3)$$



| SVIs                                                    | Formula                                                                                  | Literature    |
|---------------------------------------------------------|------------------------------------------------------------------------------------------|---------------|
| Simple ratio                                            | $SR_1 = R_{800}/R_{550}$                                                                 | 44            |
|                                                         | $SR_2 = R_{750}/R_{550}$                                                                 | 44            |
|                                                         | $SR_3 = R_{800}/R_{680}$                                                                 | In this study |
|                                                         | $SR_4 = R_{750}/R_{700}$                                                                 | In this study |
|                                                         | $SR_5 = R_{800}/R_{670}$                                                                 | 45            |
|                                                         | $SR_6 = R_{795}/R_{670}$                                                                 | In this study |
|                                                         | $SR_7 = R_{740}/R_{720}$                                                                 | In this study |
| Green indices                                           | Green I = $R_{570}/R_{670}$                                                              | 36            |
| Disease–Water Stress Index                              | $DSWI_1 = R_{800}/R_{1660}$                                                              | 46            |
|                                                         | $DSWI_2 = R_{1660}/R_{550}$                                                              | 46            |
|                                                         | $DSWI_3 = R_{1660}/R_{680}$                                                              | 46            |
|                                                         | $DSWI_4 = R_{550}/R_{680}$                                                               | 46            |
|                                                         | $DSWI_5 = (R_{800} + R_{550})/(R_{1660} + R_{680})$                                      | 46            |
| Anth reflectance index                                  | $ARI = 1/R_{550} - 1/R_{700}$                                                            | 47            |
| Moisture stress index                                   | $MSI = R_{1600}/R_{820}$                                                                 | 48            |
| Blue indices                                            | $BI = R_{450}/R_{490}$                                                                   | 49            |
| Pigment specific simple ratio chlorophyll b             | $PSSR_b = R_{800}/R_{635}$                                                               | 50            |
| Pigment specific simple ratio carotenoids               | $PSSR_c = R_{800}/R_{500}$                                                               | 50            |
| Carotenoid indices                                      | $CI_1 = R_{515}/R_{570}$                                                                 | 51            |
|                                                         | $CI_2 = R_{520}/R_{500}$                                                                 | 51            |
| Water index                                             | $WI = R_{900}/R_{970}$                                                                   | 52            |
| Red edge                                                | $RE = R_{750}/R_{710}$                                                                   | 22            |
| Vogelmann                                               | $Vog = R_{740}/R_{720}$                                                                  | 53            |
| Redness index                                           | $RI = R_{700}/R_{670}$                                                                   | 54            |
| Normalized difference vegetation index                  | $NDVI = (R_{800} - R_{670})/(R_{800} + R_{670})$                                         | 55            |
| Normalized difference                                   | $ND_1 = (R_{800} - R_{680})/(R_{800} + R_{680})$                                         | 56            |
|                                                         | $ND_2 = (R_{750} - R_{660})/(R_{750} + R_{660})$                                         | 44            |
|                                                         | $ND_3 = (R_{750} - R_{705})/(R_{750} + R_{705})$                                         | 44            |
|                                                         | $ND_4 = (R_{755} - R_{705})/(R_{755} + R_{705})$                                         | In this study |
|                                                         | $ND_5 = (R_{680} - R_{500})/R_{750}$                                                     | In this study |
| Modified ND                                             | $mND = (R_{750} - R_{445})/(R_{705} + R_{705} - R_{445})$                                | 57            |
| Modified simple ratio                                   | $mSR_1 = (R_{750} - R_{445})/(R_{705} + R_{445})$                                        | 56            |
|                                                         | $mSRI_2 = \frac{(R_{800} - R_{670}) - 1}{\sqrt{(R_{800} + R_{670}) + 1}}$                | 56            |
|                                                         | $mSRI_3 = (R_{750} - R_{445})/(R_{750} + R_{445})$                                       | In this study |
| NDWI-hyperion                                           | $NDWI-Hyp = (R_{1070} - R_{1200})/(R_{1070} + R_{1200})$                                 | 58            |
| ND water index                                          | $NDWI = (R_{860} - R_{1240})/(R_{860} + R_{1240})$                                       | 59            |
| Structure-intensive pigment index                       | $SIPI = (R_{800} - R_{445})/(R_{800} + R_{680})$                                         | 60            |
| Photochemical reflectance index                         | $PRI_1 = (R_{515} - R_{531})/(R_{515} + R_{531})$                                        | 61            |
|                                                         | $PRI_2 = (R_{534} - R_{565})/(R_{534} + R_{565})$                                        | 62            |
|                                                         | $PRI_3 = (R_{530} - R_{570})/(R_{530} + R_{570})$                                        | 63            |
| Modified chlorophyll absorption in reflectance index    | $MCARI = [(R_{700} - R_{670}) - 0.2*(R_{700} - R_{550})]/(R_{700}/R_{670})$              | 64            |
| Transformed chlorophyll absorption in reflectance index | $TCARI = 3*[(R_{700} - R_{670}) - 0.2*(R_{700} - 550)]/(R_{700}/R_{670})$                | 65            |
| Renormalized difference vegetation index                | $RDVI = \frac{(R_{800} - R_{670})}{\sqrt{(R_{800} + R_{670})}}$                          | 66            |
| Triangular veg. index                                   | $TVI = 0.5*120*(R_{750} - R_{550}) - 200*(R_{670} - R_{550})$                            | 67            |
| Average in the range of 750–850 nm                      | AVE = average between $R_{750}$ and $R_{850}$                                            | 64            |
| Modified chlorophyll-absorption-integral                | $mCAI = \frac{(R_{752} + R_{545})(R_{752} - R_{545})}{2} - \sum_{545}^{752} nm(R*1.158)$ | 68            |
| Red green ratio index                                   | $RGRI = \frac{\sum_{i=635}^{695} nm^i R}{\sum_{i=495}^{555} nm^i R}$                     | 69            |
| Red edge position                                       | $REP = 700 + 40 (R_{RE} - R_{700}) / (R_{740} - R_{700})$                                | 70            |
| Blue green pigment indices                              | $BGI_1 = R_{400}/R_{550}$                                                                | 71            |
|                                                         | $BGI_2 = R_{450}/R_{550}$                                                                | 71            |
| Blue red pigment indices                                | $BRI_1 = R_{400}/R_{690}$                                                                | 71            |
|                                                         | $BRI_2 = R_{450}/R_{690}$                                                                | 71            |
| Lichtenthaler indices                                   | $LIC = R_{440}/R_{740}$                                                                  | 72            |
| Carter Indices                                          | $CTR_1 = R_{695}/R_{420}$                                                                | 73            |
|                                                         | $CTR_2 = R_{695}/R_{760}$                                                                | 73            |
| Modified triangular vegetation index                    | $MTVI = 1.2*(1.2*(R_{800} - R_{550}) - 2.5*(R_{670} - R_{550}))$                         | 74            |

**Table 4.** List of spectral vegetation indices cited in previous studies and established in this study.



**Figure 4.** Flowchart of data processing.

$$\text{error rate} = \frac{\text{leaves } C \text{ samples misclassified into other categories}}{\text{all leaves samples in } C} \quad (4)$$

where C represents the categories: ‘healthy’, ‘asymptomatic’, ‘early stage’ and ‘late stage’.

The classification steps are as follows:

- Step 1: Randomly divide the SVIs of all of the leaf samples into a training set and a testing set according to the ratio of 70%:30%.
- Step 2: Determine the category of each sample in the training set by lazy learning, such KNN, so they are treated as known groups.
- Step 3: Randomly choose a new leaf sample in the testing set.
- Step 4: Calculate the Mahalanobis distance from the new leaf sample to each sample in the known groups.
- Step 5: According to the distance, select KNN around the new leaf sample.
- Step 6: Choose the highest frequency of 4 categories amongst KNN and determine the new sample category.
- Step 7: Repeat steps 3 to 6 until no more samples exist in the testing set.
- Step 8: Calculate the accuracy and the error rate.

## References

1. Babadoost, M. 914 edn 85–92 (International Society for Horticultural Science (ISHS), Leuven, Belgium).
2. Gleason, M. L., Gitaitis, R. D. & Ricker, M. D. Recent progress in understanding and controlling bacterial canker of tomato in Eastern North America. *Plant Disease* **77**, 1069–1076 (1993).
3. Cox, R. Control of bacterial spot of tomato in southern Florida [Xanthomonas vesicatoria]. *Plant Diseases* (1982).
4. Batuman, O., Kuo, Y. W., Palmieri, M., Rojas, M. R. & Gilbertson, R. L. Tomato chocolate spot virus, a member of a new torradovirus species that causes a necrosis-associated disease of tomato in Guatemala. *Archives of Virology* **155**, 857–869. <https://doi.org/10.1007/s00705-010-0653-9> (2010).
5. Chandra, A., Keizerwerd, A. T., Que, Y. & Grisham, M. P. Loop-mediated isothermal amplification (LAMP) based detection of Colletotrichum falcatum causing red rot in sugarcane. *Molecular biology reports*, 1–8 (2015).
6. Bukhari, K. A., Nithya, K., Valluvariparidasan, V., Paranidharan, V. & Velazhahan, R. Detection of Colletotrichum falcatum causing red rot of sugarcane by enzyme-linked immunosorbent assay. *Archives Of Phytopathology And Plant Protection* **45**, 823–830 (2012).
7. Debode, J. *et al.* Quantitative detection and monitoring of Colletotrichum acutatum in strawberry leaves using real-time PCR. *Plant Pathology* **58**, 504–514 (2009).
8. Tapia-Tussell, R. *et al.* PCR-Based Detection and Characterization of the Fungal Pathogens Colletotrichum gloeosporioides and Colletotrichum capsici Causing Anthracnose in Papaya (Carica papaya L.) in the Yucatan Peninsula. *Mol Biotechnol* **40**, 293–298. <https://doi.org/10.1007/s12033-008-9093-0> (2008).
9. Sankaran, S., Mishra, A., Ehsani, R. & Davis, C. A review of advanced techniques for detecting plant diseases. *Computers and Electronics in Agriculture* **72**, 1–13 (2010).
10. Xing, J., Symons, S., Shahin, M. & Hatcher, D. Detection of sprout damage in Canada Western Red Spring wheat with multiple wavebands using visible/near-infrared hyperspectral imaging. *biosystems engineering* **106**, 188–194 (2010).
11. Jones, C., Jones, J. & Lee, W. Diagnosis of bacterial spot of tomato using spectral signatures. *Computers and Electronics in Agriculture* **74**, 329–335 (2010).
12. Vigneau, N., Ecarnot, M., Rabatel, G. & Roumet, P. Potential of field hyperspectral imaging as a non destructive method to assess leaf nitrogen content in Wheat. *Field Crops Research* **122**, 25–31. <https://doi.org/10.1016/j.fcr.2011.02.003> (2011).
13. Stroppiana, D., Boschetti, M., Brivio, P. A. & Bocchi, S. Plant nitrogen concentration in paddy rice from field canopy hyperspectral radiometry. *Field Crops Research* **111**, 119–129. <https://doi.org/10.1016/j.fcr.2008.11.004> (2009).
14. Zhao, D., Reddy, K. R., Kakani, V. G. & Reddy, V. Nitrogen deficiency effects on plant growth, leaf photosynthesis, and hyperspectral reflectance properties of sorghum. *European Journal of Agronomy* **22**, 391–403 (2005).
15. Mahlein, A.-K., Steiner, U., Dehne, H.-W. & Oerke, E.-C. Spectral signatures of sugar beet leaves for the detection and differentiation of diseases. *Precision Agriculture* **11**, 413–431 (2010).

16. Panda, S. S., Ames, D. P. & Panigrahi, S. Application of Vegetation Indices for Agricultural Crop Yield Prediction Using Neural Network Techniques. *Remote Sensing* **2**, 673 (2010).
17. Brantley, S. T., Zinnert, J. C. & Young, D. R. Application of hyperspectral vegetation indices to detect variations in high leaf area index temperate shrub thicket canopies. *Remote Sensing of Environment* **115**, 514–523, <https://doi.org/10.1016/j.rse.2010.09.020> (2011).
18. Thenkabail, P. S., Smith, R. B. & De Pauw, E. Hyperspectral Vegetation Indices and Their Relationships with Agricultural Crop Characteristics. *Remote Sensing of Environment* **71**, 158–182, [https://doi.org/10.1016/S0034-4257\(99\)00067-X](https://doi.org/10.1016/S0034-4257(99)00067-X) (2000).
19. Devadas, R., Lamb, D., Simpfendorfer, S. & Backhouse, D. Evaluating ten spectral vegetation indices for identifying rust infection in individual wheat leaves. *Precision Agriculture* **10**, 459–470 (2009).
20. Rumpf, T. et al. Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance. *Computers and Electronics in Agriculture* **74**, 91–99, <https://doi.org/10.1016/j.compag.2010.06.009> (2010).
21. Mahlein, A. K. et al. Development of spectral indices for detecting and identifying plant diseases. *Remote Sensing of Environment* **128**, 21–30, <https://doi.org/10.1016/j.rse.2012.09.019> (2013).
22. Zarco-Tejada, P. J., Miller, J. R., Noland, T. L., Mohammed, G. H. & Sampson, P. H. Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. *Geoscience and Remote Sensing, IEEE Transactions on* **39**, 1491–1507, <https://doi.org/10.1109/36.934080> (2001).
23. Debode, J. et al. Latent entry and spread of *Colletotrichum acutatum* (species complex) in strawberry fields. *Plant Pathology* (2014).
24. Raj, M. et al. A Highly Sensitive Nested-PCR Method Using a Single Closed Tube for the Detection of *Colletotrichum gloeosporioides* causing Greater Yam Anthracnose. *Journal of Root Crops* **39**, 163–167 (2015).
25. Chen, Y. Y. et al. A quantitative real-time PCR assay for detection of *Colletotrichum lindemuthianum* in navy bean seeds. *Plant Pathology* **62**, 900–907, <https://doi.org/10.1111/j.1365-3059.2012.02692.x> (2013).
26. Horler, D., Dockray, M. & Barber, J. The red edge of plant leaf reflectance. *International Journal of Remote Sensing* **4**, 273–288 (1983).
27. Filella, I. & Penuelas, J. The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. *International Journal of Remote Sensing* **15**, 1459–1470, <https://doi.org/10.1080/01431169408954177> (1994).
28. Huber, L. & Gillespie, T. Modeling leaf wetness in relation to plant disease epidemiology. *Annual review of phytopathology* **30**, 553–577 (1992).
29. Monroe, J. S., Santini, J. B. & Latin, R. A Model Defining the Relationship Between Temperature and Leaf Wetness Duration, and Infection of Watermelon by *Colletotrichum orbiculare*. *Plant Disease* **81**, 739–742, <https://doi.org/10.1094/PDIS.1997.81.7.739> (1997).
30. Tucker, C. J. Remote sensing of leaf water content in the near infrared. *Remote Sensing of Environment* **10**, 23–32, [https://doi.org/10.1016/0034-4257\(80\)90096-6](https://doi.org/10.1016/0034-4257(80)90096-6) (1980).
31. Carter, G. A. Primary and secondary effects of water content on the spectral reflectance of leaves. *American Journal of Botany*, 916–924 (1991).
32. Chen, D., Huang, J. & Jackson, T. J. Vegetation water content estimation for corn and soybeans using spectral indices derived from MODIS near- and short-wave infrared bands. *Remote Sensing of Environment* **98**, 225–236, <https://doi.org/10.1016/j.rse.2005.07.008> (2005).
33. Martin, G. B., Bogdanove, A. J. & Sessa, G. Understanding the functions of plant disease resistance proteins. *Annual review of plant biology* **54**, 23–61 (2003).
34. Jones, J. D. & Dangl, J. L. The plant immune system. *Nature* **444**, 323–329 (2006).
35. Apan, A., Held, A., Phinn, S. & Markley, J. Detecting sugarcane ‘orange rust’ disease using EO-1 Hyperion hyperspectral imagery. *International journal of remote sensing* **25**, 489–498 (2004).
36. Calderón, R., Navas-Cortés, J. A., Lucena, C. & Zarco-Tejada, P. J. 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 (2013).
37. Huang, N. & Niu, Z. Estimating soil respiration using spectral vegetation indices and abiotic factors in irrigated and rainfed agroecosystems. *Plant and Soil* **367**, 535–550, <https://doi.org/10.1007/s11104-012-1488-9> (2013).
38. Ibrahim, Y. & Al-Saleh, M. In *III International Symposium on Tomato Diseases* 914 47–51 (2010).
39. Ghaiwat, S. N. Detection and Classification of Plant Leaf Diseases Using Image processing Techniques: A Review (2014).
40. Abdullah, N. E., Rahim, A. A., Hashim, H. & Kamal, M. M. In *Research and Development, 2007. SCORED 2007. 5th Student Conference on*. 1–6 (IEEE).
41. Lu, J. et al. Field detection of anthracnose crown rot in strawberry using spectroscopy technology. *computers and electronics in agriculture* **135**, 289–299, <https://doi.org/10.1016/j.compag.2017.01.017> (2017).
42. Lu, J., Zhou, M., Gao, Y. & Jiang, H. Using hyperspectral imaging to discriminate yellow leaf curl disease in tomato leaves. *Precision Agriculture*. <https://doi.org/10.1007/s11119-017-9524-7> (2017).
43. Hughes, D. & Salathé, M. An open access repository of images on plant health to enable the development of mobile disease diagnostics. *arXiv preprint arXiv:1511.08060* (2015).
44. Gitelson, A. & Merzlyak, M. N. Spectral Reflectance Changes Associated with Autumn Senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. Leaves. Spectral Features and Relation to Chlorophyll Estimation. *Journal of Plant Physiology* **143**, 286–292, [https://doi.org/10.1016/S0176-1617\(11\)81633-0](https://doi.org/10.1016/S0176-1617(11)81633-0) (1994).
45. Jordan, C. F. Derivation of leaf-area index from quality of light on the forest floor. *Ecology*, 663–666 (1969).
46. Apan, A., Held, A., Phinn, S. & Markley, J. Detecting sugarcane ‘orange rust’ disease using EO-1 Hyperion hyperspectral imagery. *International Journal of Remote Sensing* **25**, 489–498, <https://doi.org/10.1080/01431160310001618031> (2004).
47. Gitelson, A. A., Merzlyak, M. N. & Chivkunova, O. B. Optical properties and nondestructive estimation of anthocyanin content in plant leaves. *Photochemistry and photobiology* **74**, 38–45 (2001).
48. Hunt, E. R. & Rock, B. N. Detection of changes in leaf water content using Near and Middle-Infrared reflectances. *Remote Sensing of Environment* **30**, 43–54, [https://doi.org/10.1016/0034-4257\(89\)90046-1](https://doi.org/10.1016/0034-4257(89)90046-1) (1989).
49. Calderón, R., Navas-Cortés, J. A., Lucena, C. & Zarco-Tejada, P. J. 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> (2013).
50. Blackburn, G. A. Spectral indices for estimating photosynthetic pigment concentrations: A test using senescent tree leaves. *International Journal of Remote Sensing* **19**, 657–675, <https://doi.org/10.1080/014311698215919> (1998).
51. Zarco-Tejada, P. J., González-Dugo, V. & Berni, J. A. J. Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. *Remote Sensing of Environment* **117**, 322–337, <https://doi.org/10.1016/j.rse.2011.10.007> (2012).
52. Penuelas, J., Pinol, J., Ogaya, R. & Filella, I. Estimation of plant water concentration by the reflectance Water Index WI (R900/R970). *International Journal of Remote Sensing* **18**, 2869–2875, <https://doi.org/10.1080/014311697217396> (1997).
53. Vogelmann, J., Rock, B. & Moss, D. Red edge spectral measurements from sugar maple leaves. *TitleREMOTE SENSING* **14**, 1563–1575 (1993).
54. Gitelson, A. A. & Merzlyak, M. N. Signature Analysis of Leaf Reflectance Spectra: Algorithm Development for Remote Sensing of Chlorophyll. *Journal of Plant Physiology* **148**, 494–500, [https://doi.org/10.1016/S0176-1617\(96\)80284-7](https://doi.org/10.1016/S0176-1617(96)80284-7) (1996).
55. Rouse, J. W., Haas, R. H., Schell, J. A. & Deering, D. W. *Monitoring Vegetation Systems in the Great Plains with ERTS* (1973).

56. Sims, D. A. & Gamon, J. A. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment* **81**, 337–354, [https://doi.org/10.1016/S0034-4257\(02\)00010-X](https://doi.org/10.1016/S0034-4257(02)00010-X) (2002).
57. Sims, D. A. & Gamon, J. A. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote sensing of environment* **81**, 337–354 (2002).
58. Ustin, S. L., Roberts, D. A., Gardner, M. & Dennison, P. In *Geoscience and Remote Sensing Symposium, 2002. IGARSS '02. 2002 IEEE International*. 796–798 vol. 792 (2002).
59. Gao, B.-c NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment* **58**, 257–266, [https://doi.org/10.1016/S0034-4257\(96\)00067-3](https://doi.org/10.1016/S0034-4257(96)00067-3) (1996).
60. Pen˘Uelas, J., Filella, I., Lloret, P., Mun˘Oz, F. & Vilajeliu, M. Reflectance assessment of mite effects on apple trees. *International Journal of Remote Sensing* **16**, 2727–2733, <https://doi.org/10.1080/01431169508954588> (1995).
61. Hernandez-Clemente, R., Navarro-Cerrillo, R. M., Suarez, L., Morales, F. & Zarco-Tejada, P. J. Assessing structural effects on PRI for stress detection in conifer forests. *Remote Sensing of Environment* **115**, 2360–2375, <https://doi.org/10.1016/j.rse.2011.04.036> (2011).
62. Inch, S. A. & Ploetz, R. C. Impact of laurel wilt, caused by *Raffaelea lauricola*, on xylem function in avocado, *Persea americana*. *Forest Pathology* **42**, 239–245, <https://doi.org/10.1111/j.1439-0329.2011.00749.x> (2012).
63. Gamon, J. A., Pe˜uelas, J. & Field, C. B. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sensing of Environment* **41**, 35–44, [https://doi.org/10.1016/0034-4257\(92\)90059-S](https://doi.org/10.1016/0034-4257(92)90059-S) (1992).
64. Daughtry, C. S. T., Walthall, C. L., Kim, M. S., de Colstoun, E. B. & McMurtrey Iii, J. E. Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance. *Remote Sensing of Environment* **74**, 229–239, [https://doi.org/10.1016/S0034-4257\(00\)00113-9](https://doi.org/10.1016/S0034-4257(00)00113-9) (2000).
65. Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J. & Dextraze, L. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment* **81**, 416–426, [https://doi.org/10.1016/S0034-4257\(02\)00018-4](https://doi.org/10.1016/S0034-4257(02)00018-4) (2002).
66. Roujean, J.-L. & Breon, F.-M. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sensing of Environment* **51**, 375–384, [https://doi.org/10.1016/0034-4257\(94\)00114-3](https://doi.org/10.1016/0034-4257(94)00114-3) (1995).
67. Broge, N. H. & Leblanc, E. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment* **76**, 156–172, [https://doi.org/10.1016/S0034-4257\(00\)00197-8](https://doi.org/10.1016/S0034-4257(00)00197-8) (2001).
68. Laudien, R., Bareth, G. & Doluschitz, R. In *Proceedings of the EFITA Conference, Debrecen, Hungary*. 375–381 (2007).
69. Jackson, T. J. *et al.* Vegetation water content mapping using Landsat data derived normalized difference water index for corn and soybeans. *Remote Sensing of Environment* **92**, 475–482, <https://doi.org/10.1016/j.rse.2003.10.021> (2004).
70. Guyot, G. & Baret, F. In *Spectral Signatures of Objects in Remote Sensing*. 279 (1997).
71. Zarco-Tejada, P. J. *et al.* Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. *Remote Sensing of Environment* **99**, 271–287 (2005).
72. Lichtenthaler, H., Lang, M., Sowinska, M., Heisel, F. & Miehe, J. Detection of vegetation stress via a new high resolution fluorescence imaging system. *Journal of plant physiology* **148**, 599–612 (1996).
73. Carter, G. A. Ratios of leaf reflectances in narrow wavebands as indicators of plant stress. *Remote sensing* **15**, 697–703 (1994).
74. Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J. & Strachan, I. B. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote sensing of environment* **90**, 337–352 (2004).

## Acknowledgements

The authors wish to thank University of Florida for experiment support, the Application of Infrastructure Project of Department of science and Technology of Sichuan province (2017JY0144), Key Laboratory Open Fund (szjj2017–005), Research project of key laboratory machinery and power machinery (XIHUA UNIVERSITY), Ministry of education (JYBFX-YQ-1) and Support Program of Department of science and Technology of Sichuan province (2015NZ0055) for financial support.

## Author Contributions

J.L. and Y.S. conceived and designed the experiments; J.L., Y.S., A.I.C.M. performed the experiments; J.L. analyzed the data; R.E. and S.W. contributed materials and analysis tools; J.L. wrote the manuscript. All authors reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-21191-6>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018