2950 Niles Road, St. Joseph, MI 49085-9659, USA
269.429.0300 fax 269.429.3852 hq@asabe.org www.asabe.org

An ASABE Meeting Presentation
DOI: https://doi.org/10.13031/aim.201700076
Paper Number: 1700076

A performance comparison of RGB, NIR, and depth images in immature citrus detection using deep learning algorithms for yield prediction

Daeun Choi[1], Won Suk Lee[1], John K. Schueller[2], Reza Ehsani[3], Fritz Roka[4], Justice Diamond[1]

[1]Department of Agricultural and Biological Engineering, University of Florida, Gainesville, Florida

[2]Department of Mechanical and Aerospace Engineering, University of Florida, Lake Alfred, Florida

[3]Department of Mechanical Engineering, University of California, Merced, Merced, California

[4]Southwest Florida Research and Education Center, University of Florida, Immokalee, Florida

Written for presentation at the
2017 ASABE Annual International Meeting
Sponsored by ASABE
Spokane, Washington
July 16-19, 2017

Abstract. ***Yield forecasting is important for farm management. In this study, red, green, and blue (RGB), near-infrared (NIR), and depth sensors were implemented in an outdoor machine vision system to determine the number of immature citrus in tree canopies in a citrus grove. The main objective was to compare the performances of three image data types for citrus yield forecasting. The performance comparison was conducted with two machine vision algorithm steps: 1) circular object detection for potential fruit areas and 2) classification of citrus fruit from the background. For circular object detection, circular Hough transform was used in the RGB and NIR images. For the depth images, CHOI's Circle Estimation ('CHOICE') algorithm was developed using depth divergence and vorticity to find circular objects in the depth images. The classification process was conducted using AlexNet, a deep learning algorithm for all three image types. The implementation of a convolutional neural network allowed the machine vision algorithms to remain bias-free process during feature generation and selection. NIR images performed best with 96% true positive rate for both the circular object detection and classification. A machine vision system using this image type will produce a more objective yield prediction with a higher accuracy than other types.***

Keywords. *Citrus greening, Computer vision, Image processing, Prescription map, Yield mapping*

1. Introduction

Crop yield forecasting is an important task in farm management planning. Site-specific yield forecasting systems can perform pre-harvesting inspections of agricultural fields and give early warnings of possible adverse conditions in crops such as nutrition deficiencies and disease infections. The citrus industry in Florida is confronted with an alarming crisis due to the prevalence of citrus Huanglongbing (HLB) disease which reduces crop yields in HLB-infected trees. Site-specific yield forecasting systems can alert citrus growers to identify the locations of possible HLB-infection in groves for disease treatment. Multiple studies have been conducted to predict future citrus yield using machine vision technologies. Most of the previous studies to detect immature citrus fruit on an individual tree basis utilized commercial RGB cameras due to their practicality in outdoor field applications. Kurtulmus, Lee, & Vardar (2011) developed an immature citrus detection algorithm for color images using a circular Gabor texture analysis and 'eigenfruit' approach. They reported that the fruit detection rate was 75.3% among 166 citrus fruit. A more simple and efficient algorithm was developed by Bansal, Lee, & Satish (2013). In their study, an algorithm using fast Fourier transform (FFT) was proposed to detect immature citrus in a faster way. The overall fruit detection rate in the validation set was 82.2% among 146 fruit. In a study by Zhao, Lee, & He (2016), an immature citrus detection algorithm using color, texture and shape information of the immature fruit showed 83.4% of the correct identification and 10.7% of false positives among 308 immature citrus fruit.

Recently, various types of sensors such as near-infrared (NIR) and depth cameras have become commercially available in compact sizes, permitting an efficient adoption of sensors in outdoor imaging systems. Choi, Lee, Ehsani, Schueller, & Roka (2015) developed an immature citrus detection algorithm using coordinated color and depth sensors in Kinect. In their study, potential citrus fruit areas were detected using circular Hough transform on RGB images, and texture and 3D shape information were used in the classification of citrus fruit and background. The study reported that the algorithm showed 72.1% correct identification rate and 23.2% false positive rate.

In this study, three types of sensors: RGB, NIR, and depth, were implemented in an outdoor machine vision system to determine the number of immature citrus in tree canopies in a citrus grove. This study compared the performance of the three sensors in outdoor field situation for citrus yield forecasting. The specific objective was to compare the performance of different image types for 1) circular object detection accuracy and 2) classification accuracy using deep learning.

2. Materials and Methods

The camera used for image acquisition was Kinect for Windows v2 (Microsoft, Redmond, WA). The camera was equipped with coordinated sensors of RGB, NIR, and depth so that it can acquire images of the same field of views. Image acquisition software was developed using Kinect for Windows Software Development Kit (SDK) 2.0 provided by Microsoft in C++. A field experiment was conducted during daytime around 2:00 p.m. EDT on October 13, 2016, in an experimental citrus grove in the Plant Science Research & Education Unit (PSREU) at the University of Florida in Citra, FL. The variety of citrus was Hamlin sweet orange (*citrus sinensis*). The camera was installed on a truck, and videos were recorded of four tree rows in the citrus grove. From the recorded videos, a total of 255 images per image type. They were randomly divided into 195 training images, 30 testing images, 30 validation images. Typical examples of RGB, NIR and depth images acquired by the camera are shown in Figure 1.

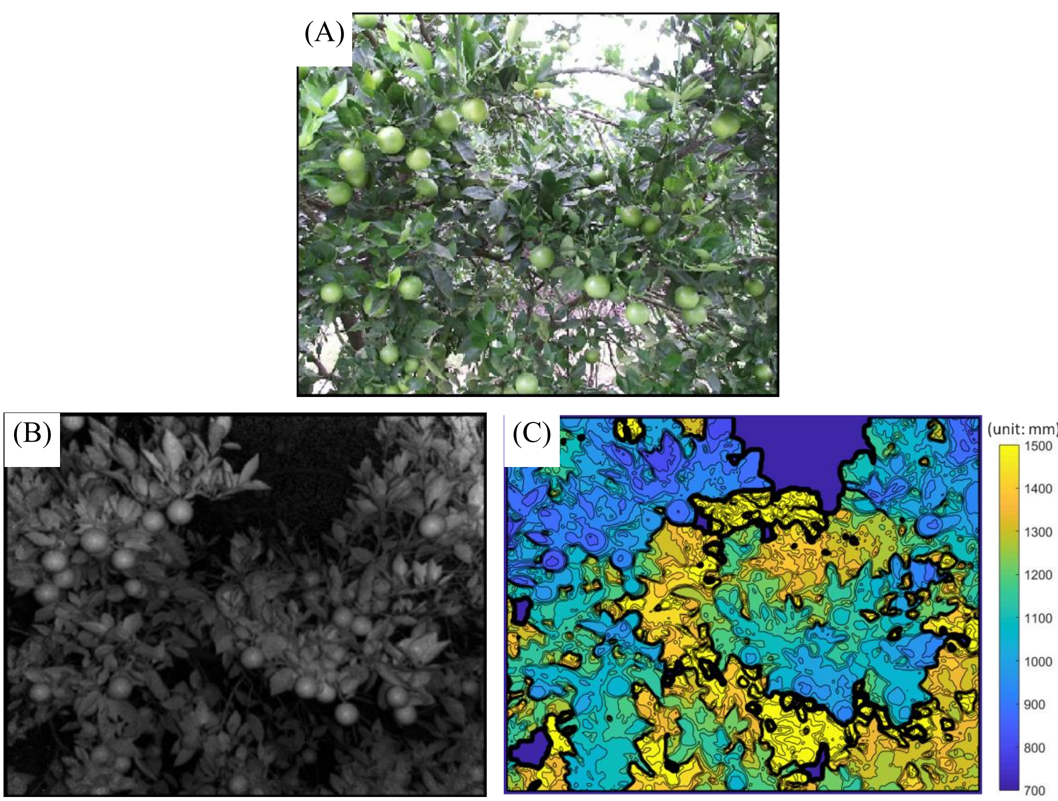

Figure 1. Example images used in this research. (A) RGB image, (B) NIR image, and (C) depth image. The RGB, NIR and depth images had the same field of views.

The resolution of the images were 1920 by 1080 pixels for color images and 512 by 424 pixels for NIR and depth images. To obtain images with the same resolution, the color images were transformed to 512 by 424 pixels using an image registration function provided in the SDK. For immature citrus detection, three machine vision algorithms (one algorithm per image type; 1. RGB, 2. NIR, and 3. depth) were developed and tested using a desktop computer (Intel(R), Core i7 3.6 GHz processor, 8 GB RAM, a 64-bit Windows 7 operating system) with a graphical processing unit (Titan X, NVIDIA, Santa Clara, CA) and MATLAB 2017a (The MathWorks Inc., Natick, Massachusetts). For a fair performance comparison, all of the three algorithms consisted of two major steps: 1) circular object detection and 2) classification of citrus fruit.

2. 1 Circular Object Detection

Regardless of citrus maturity, the fruit shape is always approximately spherical. Potential fruit areas were chosen using a circular object detection algorithm before classification of citrus fruit in the images. For two dimensional images such as the RGB and NIR, a 2D circular Hough transform (CHT) was applied to detect circular objects in the images. In depth images, images are displayed in 3D space. As a part of this work, a CHOI's circle estimation ('CHOICE') algorithm was developed to detect spherical objects in 3D space.

2.1.1 Circular Hough Transform for RGB and NIR images

When the radius of a circular object in an image is known, the CHT uses edge pixels of the image to locate the center of the circle. To obtain better edge detection results, several pre-processing techniques were applied before the CHT. In the RGB images, the intensities of R, G, and B channels were rearranged to improve the contrast of the images and reduce varying illumination conditions among images. Also, a median filter was applied to all three channels of the RGB image to reduce noise while preserving edges. The image was then converted to grayscale for the faster processing of the CHT (Figure 2A). In Figure 1B, the intensity of the NIR images was low, especially for further objects. Therefore, contrast-limited adaptive histogram equalization (CLAHE) was applied to brighten the overall intensity of the NIR (Figure 2B). An image gradient method was used for edge detection in both RGB and NIR images.

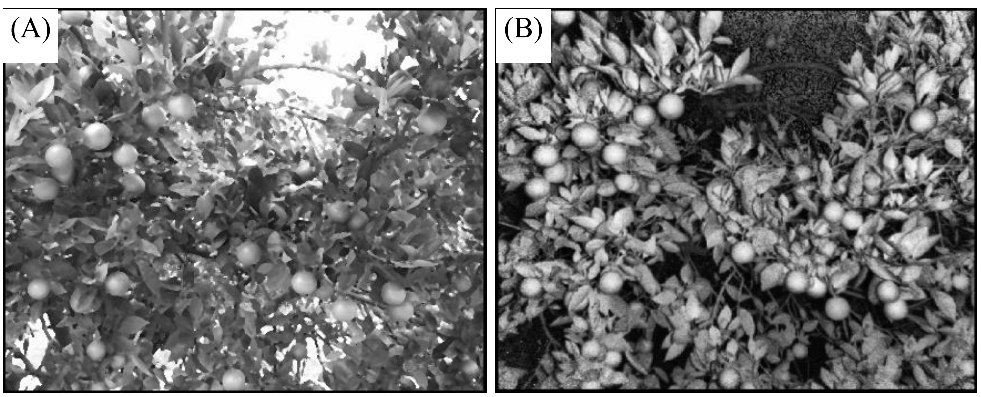

Figure 2. Preprocessing for circular Hough Transform. (A) image after preprocessing of RGB image in Figure 1A (B) image after preprocessing of NIR image in Figure 1B.

2.1.2 CHOI's Circle Estimation ('CHOICE') algorithm for depth images

In the depth images, an intensity value of each pixel represents a distance from the camera to an object surface. Therefore, the citrus fruit in the depth images exhibited a shape of convex function which had higher intensity values at the boundaries and lower values at the center of individual fruit (Figure 3A). The convexity of an object surface can be calculated using gradient vectors of depth values (Choi et al., 2015). The citrus showed a diverging pattern (Figure 3B) in a normalized gradient vector field defined in equation 1 and a clockwise circulation (Figure 3C) in a rotated gradient vector field defined in equation 2.

$$\vec{v} = \frac{\nabla D}{|\nabla D|} = \frac{1}{|\nabla D|}\left(D_x, D_y\right) \tag{1}$$

$$\vec{v}^* = \frac{1}{|\nabla D|}\left(-D_y, D_x\right) \tag{2}$$

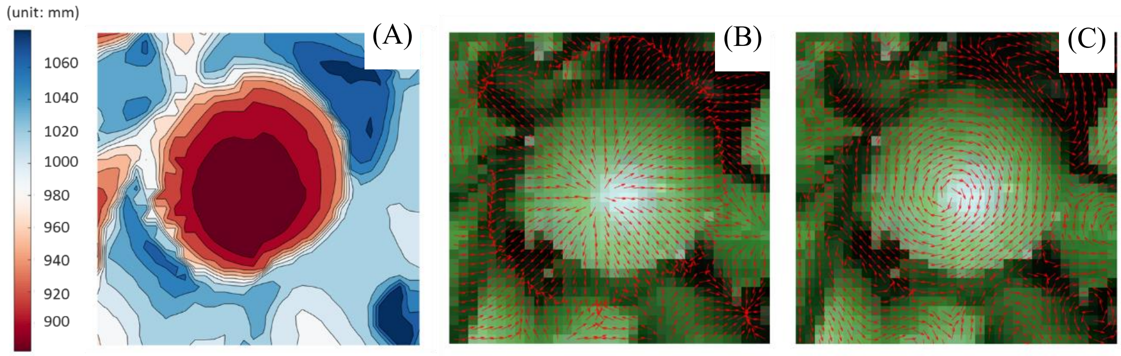

Figure 3. A depth map of a citrus fruit. (A) level sets of depth values of the citrus, (B) diverging pattern of a normalized gradient vector field, and (C) clockwise circulation of a rotated and normalized gradient vector field.

These patterns of a convex function can be quantified using divergence and vorticity formulas (Choi et al., 2015). Using Choi et al.'s method, it was found that the citrus fruit has nonnegative values inside of the fruit surface and maximum at the center in both divergence and vorticity (Figure 4). Using these characteristics, a CHOI's Circle Estimation ('CHOICE') algorithm was developed to find circular objects in the depth images. The CHOICE algorithm looked for the center of the circles that had the maximum divergence and vorticity values. Also, the boundary of the circle was chosen according to the sign of the divergence and vorticity values. Using this principle, the CHOICE algorithm found possible spherical objects in the depth images.

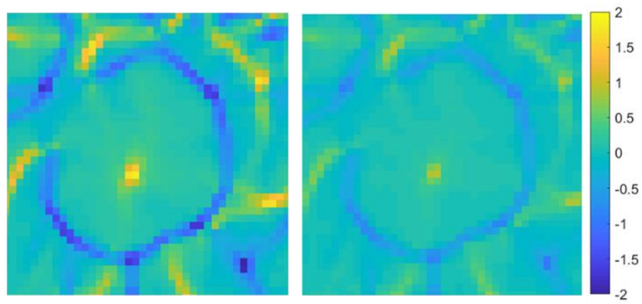

Figure 4. Vorticity and divergence calculated using the depth map in Figure 3A. (A) divergence plot of the citrus fruit surface (unit: mm/pixel), and (B) vorticity plot of the citrus fruit surface (unit: mm/pixel^2).

2. 2 Citrus Classification

After finding the potential fruit objects in the RGB, NIR, and depth images, classification using AlexNet (Krizhevsky, Sutskever, & Hinton, 2012) was conducted to detect the fruit from the background. The AlexNet is a type of deep convolutional neural net (CNN) which has a total of eight learnable layers (five convolutional and three fully connected layers). The implementation of the CNN allowed the machine vision algorithms to select the most powerful features depending on the types of images. In this way, the algorithms remained in a bias-free process from feature generation and selection. A transfer learning method was used to train the AlexNet for the three data types, RGB, NIR, and depth, separately. Since the original AlexNet was designed to classify 1000 classes, the last fully connected layer was modified to have two classes, i.e., citrus and background. A total of 2000 citrus and 2000 background samples per data type were manually cropped from 195 training images and used for training the AlexNet models. Also, 400 citrus and 400 background samples were extracted from the 30 images to be used in the testing set. For validation, 30 images per data type containing 492 citrus fruit were used. To compare the performance of the developed machine vision algorithms, the results were analyzed separately in 1) the circular object detection, and 2) the classification of the citrus using the percentages of true positives and false positives.

3. Results and Discussions

Example images showing final results are shown in Figure 5. Figure 5A, 5C, and 5E display example result images from the circular object detection using the CHT and CHOICE on RGB, NIR, and depth images, respectively. In the images, most of the citrus was detected with a small number of missed fruit. However, a considerable amount of background objects were also detected. The detected background objects were removed in the classification process as shown in Figure 5B, 5D, and 5F for the RGB, NIR and depth images, respectively. For the depth images, the results were displayed on the NIR images for better recognition of objects in the images.

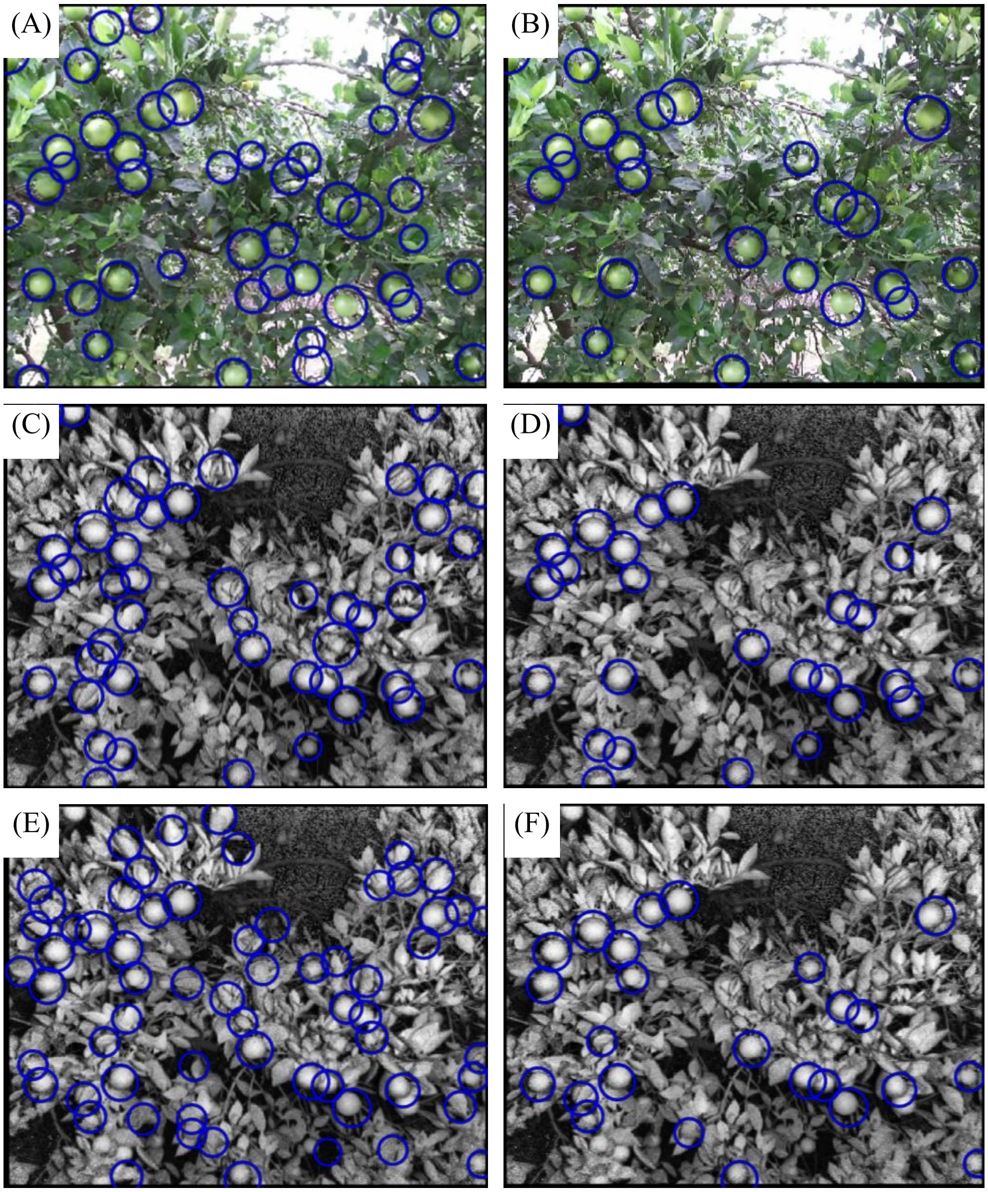

Figure 5. Example images of the results from two major steps, 1) circular object detection and 2) classification of the citrus. (A) a result of CHT on an RGB image, (B) a result of the classification on the RGB image, (C) a result of CHT on an NIR image, (D) a result of the classification on the NIR image, (E) a result of CHOICE on a depth image, and (F) a result of the classification on the depth image.

The results of the circular object detection are shown in Table 1. A volunteer conducted a validation to count actual numbers of citrus fruit, false negatives, and false positives in the RGB, NIR, and depth images. Among the three types of images, the NIR images had the best performance in the circular object detection by having 96% true positive rate and a ratio of 1:1.1 between true positives and false positives. The main reason for the NIR image's good performance was that the NIR images had a bigger intensity gap between object boundaries and other pixels, which made the edge detection process easier. Edge detection in the RGB images was relatively more difficult than in the NIR images, due to smaller intensity gaps between the boundary pixels and other pixels. Also, the depth map contained noise from the sunlight which restricted the performance of the CHOICE algorithm.

Table 1. Result of circular object detection.

	Actual Number of Fruit	True Positives (%)	False Negatives (%)	True Positives : False Positives
RGB	492	429 (87.2)	63 (12.8)	1:3.2
NIR	492	472 (96.0)	20 (4.0)	1:1.1
Depth	492	414 (84.1)	78 (15.9)	1:3.2

Table 2 shows the classification results of the RGB, NIR, and depth images for immature fruit detection. To analyze the classification results, missed fruit in the circular object detection were ignored, and only the correctly detected citrus fruit in the first step was considered. For the classification, all three image types showed acceptable accuracies. Similar to the result in Table 1, the NIR images showed the best classification performance. However, it is important to note that the RGB and depth images had higher false positives from the circular object detection process. Consequently, the RGB and depth images had higher false positive rates in the classification compared to the NIR images. Considering the results in Table 1 and 2, using the NIR image algorithm showed the best performance for immature citrus fruit detection. Also, these tables show that the final accuracy of immature citrus detection was highly affected by the performance of the circular object detection.

Table 2. The result of citrus classification.

	Actual Number of Fruit	True Positives (%)	False Negatives (%)	False Positives (%)
RGB	429	393 (91.6)	36 (8.4)	55 (12.8)
NIR	472	453 (96.0)	19 (4.0)	36 (7.6)
Depth	414	375 (90.6)	39 (9.4)	82 (19.8)

4. Conclusions

This study was conducted to analyze the performance of RGB, NIR and depth images for immature citrus fruit detection using machine vision techniques. For each image type, a machine vision algorithm was developed to include two major steps; 1) circular object detection to find potential fruit areas and 2) classification of citrus fruit from the background. Also, a new method ('CHOICE') to detect circular objects in the depth images was proposed using divergence and vorticity in a gradient vector field of depth images. In the analysis, NIR images showed the best performance for both circular object detection and classification of the citrus. The results of this study can play a significant role in evaluating the most efficient image type in an outdoor machine vision system for citrus yield prediction. Also, a machine vision system using the most efficient image type will produce a higher detection accuracy to deliver more objective and detailed yield prediction. For future work, the CHOICE algorithm can be improved to have better detection accuracy by adopting a tolerance factor to ignore noise from sunlight. Also, the developed algorithm can be implemented in real-time and combined with unmanned aerial or ground vehicles to have a fully automated citrus yield prediction system.

Bansal, R., Lee, W. S., & Satish, S. (2013). Green citrus detection using fast Fourier transform (FFT) leakage. *Precision Agriculture, 14*(1), 59-70. doi:10.1007/s11119-012-9292-3

Choi, D., Lee, W. S., Ehsani, R., Schueller, J. K., & Roka, F. (2015). *Machine vision system for early yield estimation of citrus in a site-specific manner.* Paper presented at the 2015 ASABE Annual International Meeting.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). *Imagenet classification with deep convolutional neural networks.* Paper presented at the Advances in neural information processing systems.

Kurtulmus, F., Lee, W. S., & Vardar, A. (2011). Green citrus detection using 'eigenfruit', color and circular Gabor texture features under natural outdoor conditions. *Computers and Electronics in Agriculture, 78*(2), 140-149. doi:10.1016/j.compag.2011.07.001

Zhao, C., Lee, W. S., & He, D. (2016). Immature green citrus detection based on colour feature and sum of absolute transformed difference (SATD) using colour images in the citrus grove. *Computers and Electronics in Agriculture, 124*, 243-253.