

## EVALUATING THE PERFORMANCE OF A SEMI-AUTOMATIC APPLE FRUIT DETECTION IN A HIGH-DENSITY ORCHARD SYSTEM USING LOW-COST DIGITAL RGB IMAGING SENSOR

### *Avaliação do Desempenho de Detecção Semi-Automática de Frutos de Maçã em um Pomar de Alta Densidade usando Sensor de Imagem Digital RGB de Baixo Custo*

Leonardo Josué Biffi<sup>1</sup> - ORCID: 0000-0003-3334-4507

Edson Aparecido Mitishita<sup>1</sup> - ORCID: 0000-0003-1717-7657

Veraldo Liesenberg<sup>2</sup> - ORCID: 0000-0003-0564-7818

Jorge Antonio Silva Centeno<sup>1</sup> - ORCID: 0000-0002-2669-7147

Marcos Benedito Schimalski<sup>2</sup> - ORCID: 0000-0001-7401-3881

Leo Rufato<sup>3</sup> - ORCID: 0000-0001-9545-7035

<sup>1</sup> Universidade Federal do Paraná, Programa de Pós-Graduação em Ciências Geodésicas, Curitiba – Paraná, Brasil.  
E-mails: leonardo.biffi@udesc.br, mitishita@ufpr.br, centeno@ufpr.br

<sup>2</sup> Universidade do Estado de Santa Catarina, Departamento de Engenharia Florestal, Lages – Santa Catarina, Brasil.  
E-mails: veraldo.liesenberg@udesc.br, marcos.schimalski@udesc.br

<sup>3</sup> Universidade do Estado de Santa Catarina, Departamento de Agronomia, Lages – Santa Catarina, Brasil.  
E-mail: leo.rufato@udesc.br

Received in 23<sup>rd</sup> January 2020.

Accepted in 24<sup>th</sup> April 2021.

#### Abstract:

This study investigates the potential use of close-range and low-cost terrestrial RGB imaging sensor for fruit detection in a high-density apple orchard of Fuji Suprema apple fruits (*Malus domestica* Borkh). The study area is a typical orchard located in a small holder farm in Santa Catarina's Southern plateau (Brazil). Small holder farms in that state are responsible for more than 50% of Brazil's apple fruit production. Traditional digital image processing approaches such as RGB color space conversion (e.g., *rgb*, *HSV*, *CIE L\*a\*b\**, *OHTA*[ $I_1, I_2, I_3$ ]) were applied over several terrestrial RGB images to highlight information presented in the original dataset. Band combinations (e.g., *rgb-r*, *HSV-h*, *Lab-a*,  $I''_2, I''_3$ ) were also generated as additional parameters (*C1*, *C2* and *C3*) for the fruit detection. After, optimal image binarization and segmentation, parameters were chosen to detect the fruits efficiently and the results were compared to both visual and in-situ fruit counting. Results show that some bands and combinations allowed hits above 75%, of which the following variables stood out as good predictors: *rgb-r*, *Lab-a*,  $I''_2, I''_3$  and the combinations *C2* and *C3*. The best band combination resulted from the use of *Lab-a* band and have identical results of commission, omission, and accuracy, being 5%, 25% and 75%, respectively. Fruit detection rate for *Lab-a* showed

**How to cite this article:** BIFFI, L.J.; MITISHITA, E.A.; LIESENBERG, V.; CENTENO, J.A.S.; SCHIMALSKI, M.B.; RUFATO, L. Evaluating the performance of a semi-automatic apple fruit detection in a high-density orchard system using low-cost digital rgb imaging sensor. *Bulletin of Geodetic Sciences*. 27(2): e2021014, 2021.



This content is licensed under a Creative Commons Attribution 4.0 International License.

a 0.73 coefficient of determination ( $R^2$ ), and fruit recognition accuracy rate showed 0.96  $R^2$ . The proposed approach provides results with great applicability for small holder farms and may support local harvest prediction.

**Keywords:** *Malus domestica* Borkh, fruit detection, color space, precision fructiculture, precision agriculture.

## 1. Introduction

Agriculture is crucial for the Brazilian trade balance. Improvement in management practices helped productivity grow for different cultures and per area. The agricultural sector also faced automation aiming to efficiently use resources and reduce costs. Remote sensing images acquired on orbital, aerial, and terrestrial levels can support the monitoring of a given culture at different growth stages and deliver results at different temporal and spatial scales.

Close-range data collection became a subject of study with the popularization of portable sensors, as successfully shown in Koenig et al. (2015) or Vázquez-Arellano et al. (2016). Some studies focused on the use of portable sensors in fruit growth monitoring using either active (Berk et al., 2016; Colaço et al., 2017; Escolà et al., 2017) or passive remote sensing sensors (Coelho Filho et al., 2005; Zhou et al., 2012; Linker, 2017). Given their high cost, these systems are usually not affordable for small holder farms and production estimates in several orchards are still based on visual counting techniques. This opens new perspectives for the benefits of using low-cost cameras in robotic vision (Gongal et al., 2016 and An et al., 2017), fruit detection using digital image processing techniques (Font et al., 2014a; Wei et al., 2014 and Liu et al., 2016), automatic harvesting (Font et al., 2014b and Kang and Chen, 2019) and yield estimation (Bargoti and Underwood, 2017; Dorj, Lee and Yun, 2017).

Fast and straightforward computational techniques such as the use of Red-Green-Blue (RGB) images and the transformation of these images into other color spaces to identify targets of interest in fruit growing was the subject of several studies (Behroozi-Khazaei and Maleki, 2017; Tao and Zhou, 2017; Aquino et al., 2018). For Zhou et al. (2012), the detection of the fruit before harvesting is a valuable support for better harvesting management, labor contracts, management of storage capacity, and future market planning. Fruit detection usually employs algorithms that transform the original RGB data into the Hue-Saturation-Intensity (HSI) color space and perform the segmentation of the fruits in the images by thresholding. The reported results show a high success rate, achieving a coefficient of determination ( $R^2$ ) of 0.85 when the authors compare the automatic method and the reference data obtained by visual counting.

Other algorithms based on color transformations were also evaluated by Jidong et al. (2016) in the identification of occluded fruit apples. They transformed the RGB images into the XYZ, CIE (International Commission on Illumination)  $L^*a^*b^*$  and OHTA color spaces (components I1, I2 and I3) (Ohta, Kanade and Sakai, 1980), and then tried a fixed threshold method and another of dynamic threshold, known as the OTSU method (Otsu, 1979), to segment the fruits in the images. The best results were obtained using I2, I3 and  $a^*$  color components and the dynamic segmentation method. A similar experiment was performed by Wei et al. (2014) to detect different fruits with color highlights in relation to the surrounding environment and the background. The RGB images were transformed into the OHTA color space, and a dynamic threshold segmentation proposed by OTSU was chosen. The results were satisfactory for fruit detection and segmentation, even within a complex agricultural environment. The proposal was a solution for a robot vision device for fruit recognition and reached 95% of correct recognition rates. All cited methods were reported to be of low computational demand and practical usage for remote sensing users and enthusiasts on digital image processing.

However, recent advances in low-cost RGB datasets analyses have also been reported from the transformation of the 2D information to 3D space. Data from active sensors and depth cameras (RGB-D) are thus evaluated with machine learning techniques such as deep learning (e.g. R-CNN, YOLO, SSD) for both segmentation and fruit detection (Gené-Mola et al., 2019; Koirala et al., 2019) with very promising results. The use of convolutional neural networks

in 2D images for fruit detection can also be applied (Häni, Roy and Isler, 2020 and Liu et al., 2019). Although these techniques are reported to be highly accurate in detecting objects, they also demand great computational capacity to analyze the data for training, as well as large datasets of previously sampled and labeled. Consequently, these techniques are also impracticable for small holder farms, if this information is not available. Therefore, experiments involving the potential of close-range and low-cost terrestrial RGB images employing fast and near real-time digital image processing applied for fruit detection are still necessary.

The southern region of Brazil, specifically the state of Santa Catarina, is the largest apple-producing area in the country (IBGE, 2018). More than 50% of the national apple production comes from small holder farms who have orchards smaller than ten hectares (Bittencourt et al., 2011). Despite this condition, few scientific and technological developments for the automatic detection and counting of apple fruits have been conducted so far. This is especially true for small holder farms where either expensive devices or advanced computer settings for fruit counting and production estimates are not affordable.

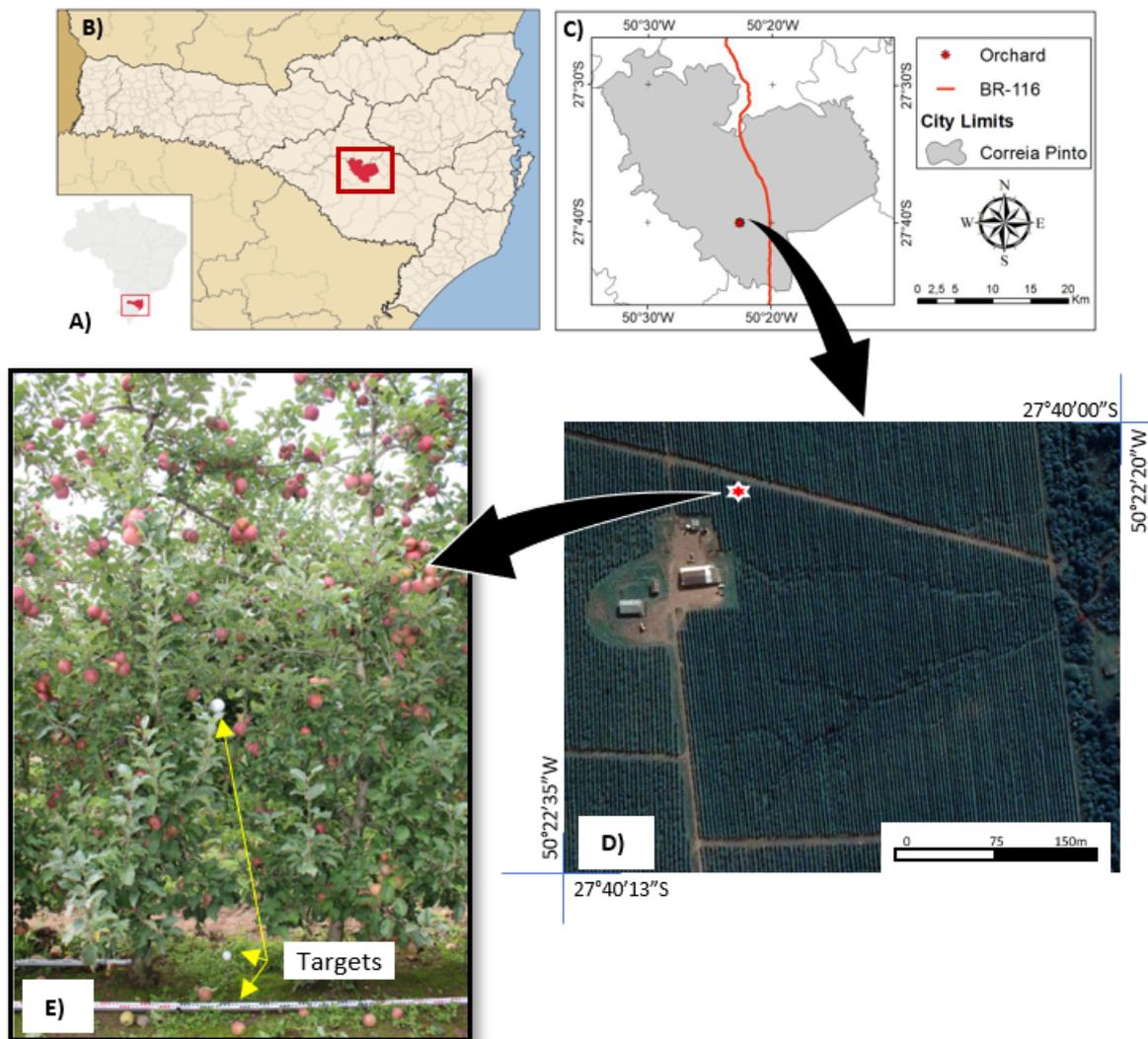
Based on this deficiency, this article presents the results of a study aimed at detecting and counting the occurrences of apples before harvesting in a commercial apple orchard with natural background, using images obtained with a low-cost portable digital camera and exploring different color spaces using standard digital image processing techniques. Such low-cost approaches that are still necessary for small-holder farmers. The research process can be summarized in three key steps: (1) digital processing of RGB image with a simple algorithm for easy use; (2) automatic fruit detection and counting; and (3) determination of apple fruit detection accuracy in close-range and low-cost terrestrial RGB imaging sensor.

## 2. Material and Methods

### 2.1 Study Area

The experiment was conducted in a typical commercial apple orchard, located in Correia Pinto city, Santa Catarina (27°40'01"S, 50°22'32"W), Brazil (Figure 1 A, B and C). The plant conduction system is designed to assume an architecture as needed by the producer. The conduction is done by pruning and a training system to support the plants. The farm deploys an orchard with two different cultivars to promote pollination. The analyzed area has two planted cultivars, Fuji Suprema and Gala, planted in a configuration of 4 × 2 rows. In summary, four Fuji rows are always followed by two Gala rows. However, for this research we only selected the Fuji Suprema cultivar. This ratio is the same across the whole orchard. Planting density is very high (Petri et al., 2011), with a 0.8 × 3.5 m spacing what resulted in a density of 3,570 plants per ha. The planting rows follow N-S orientation for better exposure to the sun (Figure 1D).

Field surveys and experiments were performed weekly starting from September 2018, and encompassing therefore different plant growth stages. The fruits of this cultivar acquire the typical red color since the early stage of fruit development. The plants were in full production in March 2019, two weeks before harvesting. At this stage, the plants increased the maximum leaf area index (LAI) forming a green wall, and the fruits were already ripe and had complete development, which allows a better detection and consequently a more realistic production estimate (Cheng et al., 2017). This study aims only to detect fruits and therefore any inference physiological status or quality of the fruits are beyond the scopus of this study.

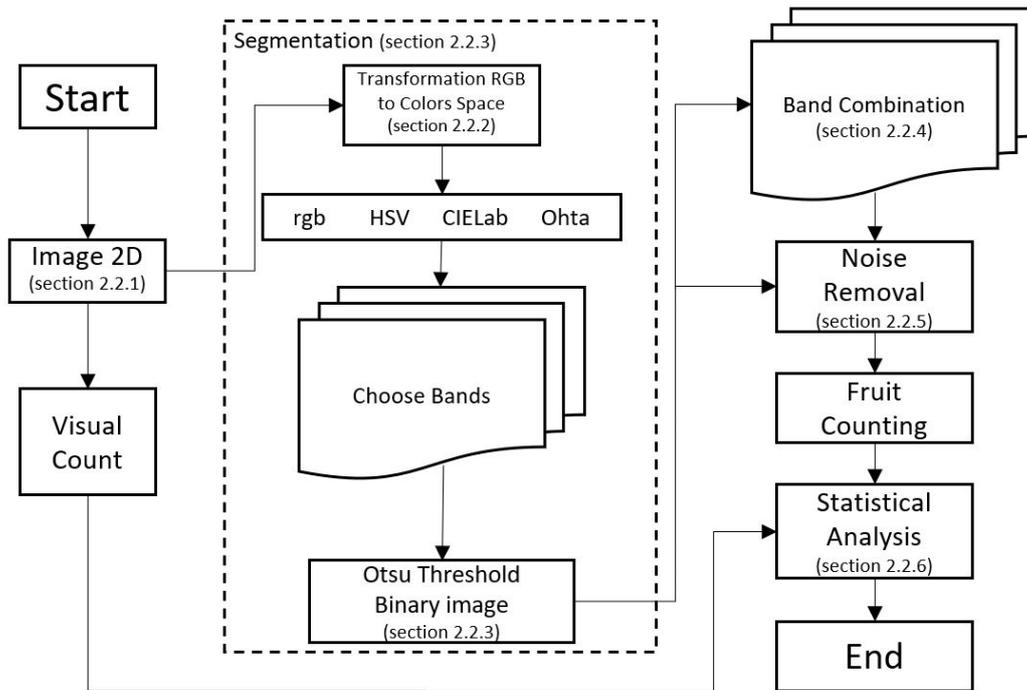


**Figure 1:** Location of the study area: A) Brazil; B) State of Santa Catarina – SC; C) Municipality boundary of Correia Pinto and Orchard Site; D) Aerial view of the Orchard and Experiment Site in true color composition; E) Example of a terrestrial RGB images used for fruit detection.

Although the apple fruits can be visually identified by its typical red color after the fruit set phase, the leaves still cause occlusions, mainly when the fruits are very small. During the development of the apple fruit, the load of fruits per plant tends to decrease due to thinning practices and natural fall. Therefore, fruit counting procedures executed close to the harvesting period provide a better estimate of the fruit production, and was then selected for this study.

## 2.2 Method of fruit detection

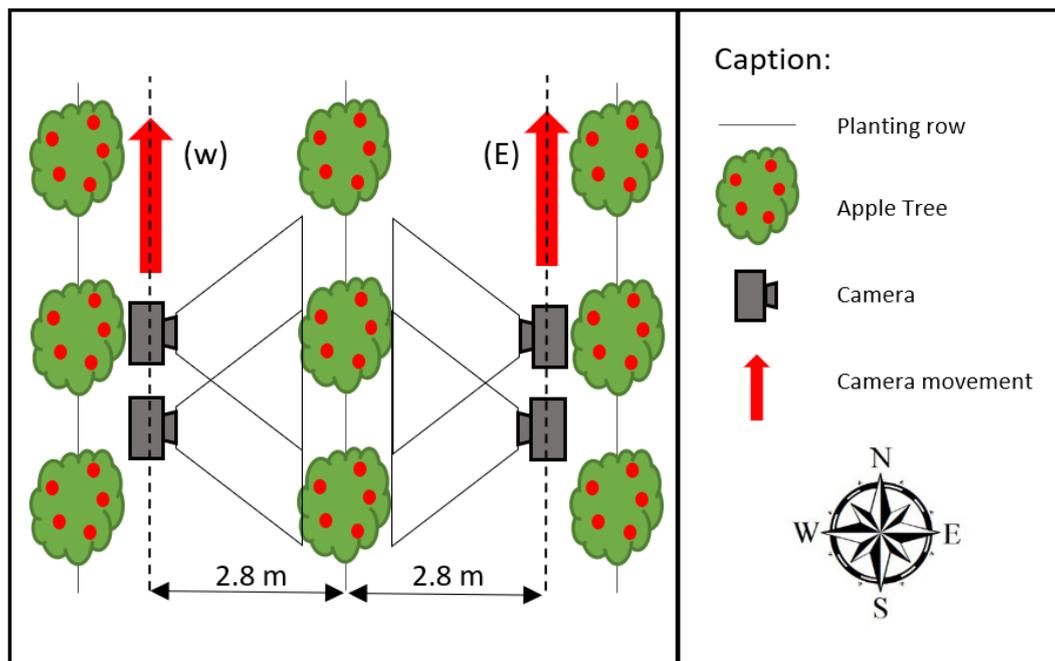
The proposed methodology for apple counting is presented in Figure 2. All algorithms were developed using the Matlab© environment installed on a notebook with Intel Core i7- 6700H, 2.60 GHz, 16 GB memory, and GPU NVIDIA GTX 960M with 2GB graphics memory. The methodology can be divided in the following major steps: 1) image acquisition; 2) color space transformation; 3) segmentation; 4) band combination; 5) noise removal and 6) statistical analysis. The details of the used methods are introduced and better explained in subsections 2.2.1 to 2.2.6.



**Figure 2:** Methodological flowchart of the proposed low-cost fruit detection approach.

### 2.2.1 Image acquisition

The process starts with the acquisition of close-range terrestrial images from both sides of the plants. A digital camera (Canon® EOS Rebel – T6, Tokyo, Japan) was used with a complementary metal-oxide semiconductor (CMOS) sensor of 5184×3456 pixels (17.9 Megapixels) and 4.3 μm (Micrometers =  $1 \times 10^{-6}$  m) pixel size was used. The nominal focal length used was 18 mm. The image capture sequence was designed to provide the maximal coverage of the plant under two constrains: to use only one shot per position and that the image is acquired perpendicular to the alignment of the trees. The distance between stations was set according to Liang et al., (2014) in “Stop and go” mode. After one image is obtained, a small displacement is made parallel to the direction of the orchard planting rows. A series of consecutive images was captured, simulating a sensor fixed on a moving vehicle. Ten Images were captured on each side of the rows, resulting in 20 images that were selected for this study. Although the number of images is low and considering that it is conducted only in one orchard, the viability of a proposed low-cost approach is still valid. Figure 3 summarizes the schematic strategy employed in the field measurements. The camera was operated manually at the operator’s face height, and the images were taken with the camera in portrait position, at a distance of nearby 2.8 m from the planting row. The high density orchard and plants have an average height of 3.5 m. It is important to mention that the camera’s field of view was sometimes not sufficient to imagine the top of all plants. However plants smaller than 3.0 m, pruned or with anti-hail coverage could be completely imaged. The day was cloudy, and the camera’s auxiliary flash was active. Acquisitions were made in the date of 29 the March 2019, between 12 p.m. and 2 p.m. local time to avoid shadows from neighboring plants. It allows to explore the exposure of the apple trees and the fruits as shown in Table 1 and Figure 4. It is also important to mention that no artificial background was adopted, so that the proposed approach consist in natural field conditions.



**Figure 3:** Top view of image acquisition and camera movement.

Figure 1E is an example of one of the images obtained in this study; it is an image of the West side of the row and it shows that it is not possible to separate the individual trees in the image due to the distance of 0.8 m between the plants along the row. This is the reason why they are also called “green wall”. It is also observed that there are fruits on the plants and on the ground. Those on the ground fell naturally. Stadia marks and white spheres of different sizes were used to fix the image scale (Figure 1E).

**Table 1:** Sun elevation and Sun Azimuth angle values at different times for March 29, 2019.

Local time	Sun Elevation (degrees)	Sun Azimuth (degrees)
6:29 a.m. (sunrise)	-0.83	86.67
8:00 a.m.	18.9	75.55
10:00 a.m.	43.03	54.49
12:00 p.m.	58.27	12.57
2:00 p.m.	51.55	320.35
4:00 p.m.	29.9	292.38
6:00 p.m.	4.16	276.18
6:22 p.m. (sunset)	-0.83	273.55

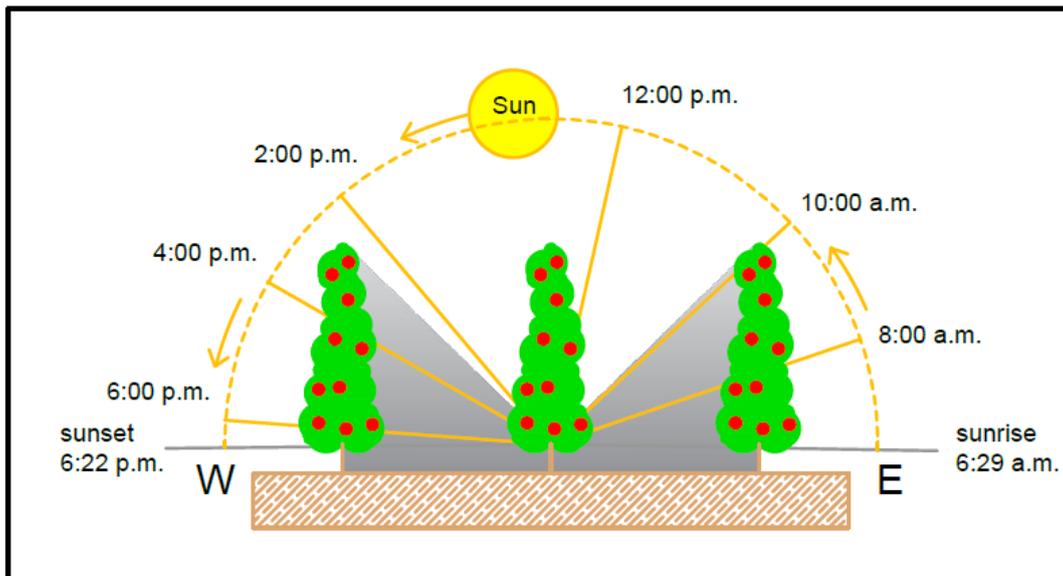


Figure 4: Artistic projection of shadows of the neighbor planting rows on March 29, 2019.

### 2.2.2 Color space transformation

The original RGB images were then transformed into several spaces (e.g., HSV, OHTA,  $L^*a^*b^*$  and  $rgb$ ). To evaluate the use of different color systems and chose the best color system for fruit apple detection.

The first option was to normalize the RGB values according to Equation 1. This transformation allows separating the color information from the intensity (Shaik et al., 2015):

$$\begin{cases} r = \frac{R}{R + G + B} \\ g = \frac{G}{R + G + B} \\ b = \frac{B}{R + G + B} \end{cases} \quad (1)$$

where:  $r$ ,  $g$  e  $b$  are the normalized red, green, and blue color values, respectively. And the  $R$ ,  $G$  and  $B$  values are the original image values of red, green and blue, respectively.

Another commonly used system is HSV that stands for the components Hue, Saturation, and Value. It is also known as HSB (Hue, Saturation and Brightness) and is defined as follows (Dorj, Lee and Yun, 2017):

- i) Hue: is the measure of the average wavelength of light that the object reflects or emits. The values range from  $0^\circ$  to  $360^\circ$ , representing all visible colors and can be normalized to 0 to 1.
- ii) Saturation: can be called the “purity” of color. Lower values are found at grey shades. The higher the saturation, the “purer” is the color. It varies between 0 (white/grey) to 1 (full saturation).
- iii) Value: stands for the brightness of the color and varies between from 0 (black) and 1 (white).

The system is the result of a nonlinear transform, as described in equations 2-4 (Shaik et al., 2015):

$$H = \arccos \left( \frac{\frac{1}{2}(2R - G - B)}{\sqrt{(R - G)^2 - (R - B)(G - B)}} \right) \quad (2)$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \quad (3)$$

$$V = \max(R, G, B) \quad (4)$$

For the transformation from RGB color space to CIE space  $L^*a^*b^*$  the formulas presented by Jidong et al. (2016) were used:

$$\begin{cases} L = 0.2126R + 0.7152G + 0.0722B \\ a = 1,4749(0.2213R - 0.339G + 0.1177B) + 128 \\ b = 0.6245(0.1949R + 0.6057G - 0.8006B) + 128 \end{cases} \quad (5)$$

The orthogonal color components in the OHTA color space ( $I_1, I_2, I_3$ ) are obtained applying linear transformation to the R, G and B the components are independent, as displayed in equation 6 (Ohta, Kanade and Sakai, 1980):

$$\begin{cases} I_1 = (R + G + B)/3 \\ I_2 = (R - B)/2 \\ I_3 = (2G - R - B)/4 \end{cases} \quad (6)$$

One can notice that in equation 6 is equivalent to the intensity or brightness. The other two components define a plane that is perpendicular to the first component. There are several possible vectors to represent this plane and one can chose a pair that better enhances a certain color. To highlight red apple fruits, the vectors in equation (7) were adopted (adapted from Ohta, Kanade and Sakai, 1980).

$$\begin{cases} I''_2 = (R - G) \\ I''_3 = (2R - G - B) \end{cases} \quad (7)$$

### 2.2.3 Image segmentation and binarization

Image segmentation is a computational process that allows dividing the image into homogeneous regions, highlighting areas of interest in the image. There are several solutions to perform this task. Among them there is the thresholding method. Thresholding consists in identifying the gray level value that enables separating the object of interest from the background. As even images of the same area may be obtained under different lighting conditions, the threshold may vary from image to image, which is not desired (Jidong *et al.*, 2016). A possible solution to this problem is to use a dynamic threshold method, where the operator does not influence the choice of threshold.

The thresholding method proposed by Otsu (1979) computes the optimal threshold " $T$ " by analyzing the histogram of the image and selecting the best gray level that would produce two clusters: one with low values (background) and one with high values, which allows binarizing the image.

Considering that  $T$  may assume any possible gray level values in the image, all possible values are evaluated based on probabilities computed for the resulting clusters:  $w_0, \mu_0, w_1, \mu_1$  and  $\sigma_B^2$ . First, for each possible threshold it is computed: the average of the gray values in the foreground and in the background ( $\mu_0, \mu_1$ ) and the global mean ( $\mu_T$ ).

Then, probability functions are computed for each cluster:

- i)  $w_0$  the probability of the foreground cluster, given by the ratio of the number of foreground pixels to the total number of pixels.
- ii)  $w_1$  is the probability of the background, given by the ratio of the number of background pixels to the total number of pixels.

The optimal threshold value  $T$  will be the one that maximizes the variance between classes ( $\sigma_B^2$ ), according to Equations 8-10 (Otsu, 1979).

$$\mu_T = w_0 \times \mu_0 + w_1 \times \mu_1 \quad (8)$$

$$\sigma_B^2 = w_0 \times (\mu_0 - \mu_T)^2 + w_1 \times (\mu_1 - \mu_T)^2 \quad (9)$$

$$\sigma_B^2 = w_0 \times w_1 \times (\mu_0 - \mu_1)^2 \quad (10)$$

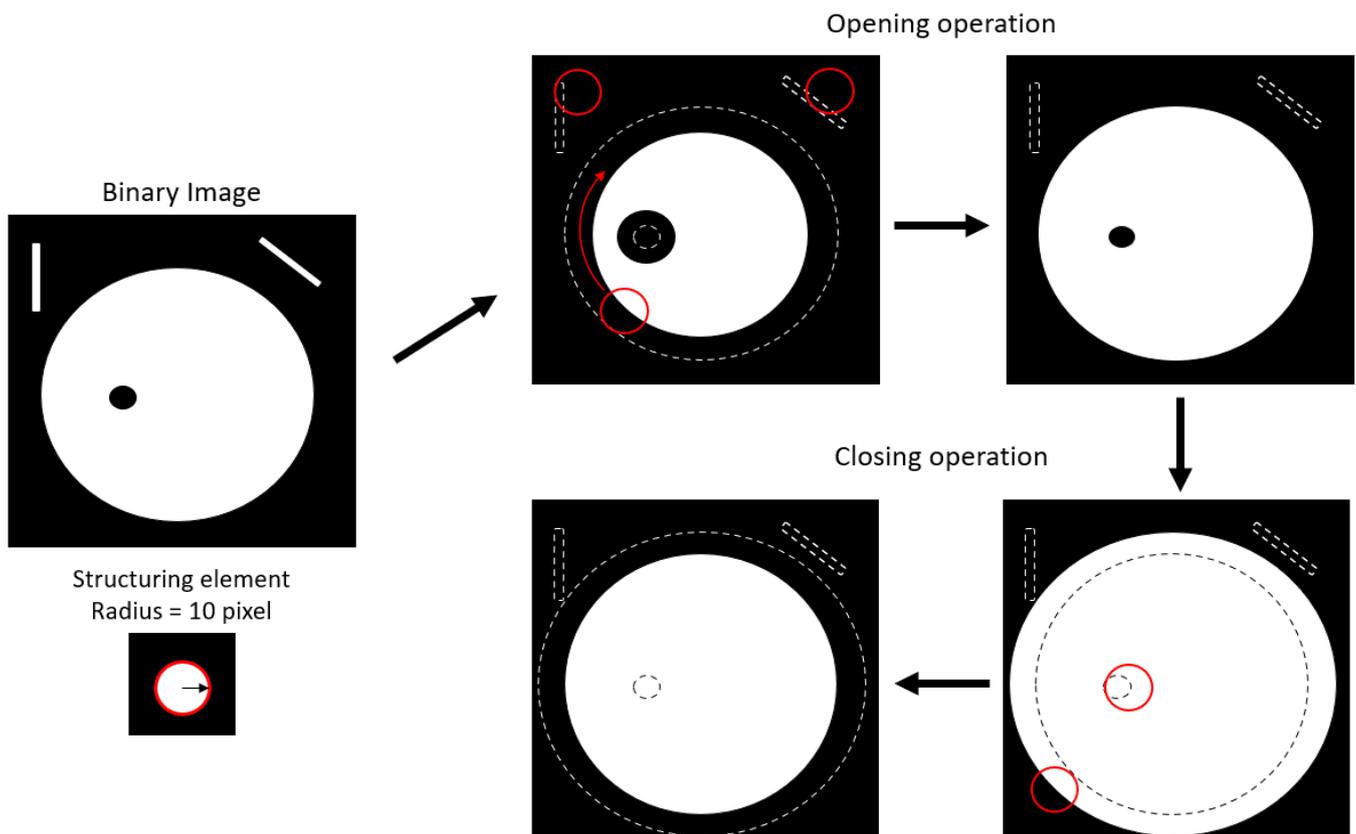
### 2.2.4 Band combination

This work presents a proposal to combine the images that had better responses for apple fruit detection after the segmentation step. The main idea behind this procedure is to select the images from the previous step and combine them in an additive way. As the images are binary, they have only values 0 and 1, corresponding to the background identification and fruit, respectively. Thus, combining two images (for example) with areas that correspond to the same detected fruit would result in a new image with a value of 2 in this area of the fruit.

The corresponding background areas in both images remain with value 0 and matching areas that differ in values in both images will assume value 1. Thus, a simple threshold ( $T$ ) of  $n-1 < T < n$ , where  $n$  is the number of images used in the combination, would assist in the removal of discrepant noises originating from the bands used.

### 2.2.5 Removal of noise from images

After image binarization small regions and blanks were removed applying mathematical morphological operations, like opening and closing, using as structuring element a 10-pixel radius circle (Figure 5). The most common errors are caused by twig edges, leaf tips, small fruits from the rows in the background, and small visible fruit regions.



**Figure 5:** Binary image noise removal scheme, applying opening and closing mathematical morphological operation, using circle structuring element.

### 2.2.6 Statistical analysis

The results of the fruit detection process of each image were used, matching the visual fruit count performed in the laboratory on each RGB image. Therefore, each single apple fruit received a location mark on the RGB image. In summary, all visible or partially visible apple fruits on the plant as well as fruits on the ground floor were marked.

The detected fruits that coincided with the visual markings were counted as true positive (TP). On opposite cases, this hit was considered a false positive (FP). All remaining fruits that were not detected by the algorithm were counted as false negative (FN).

Two-dimensional plots were then generated considering the existing relation between the number of objects detected vs. the number of apples from visual counting. Similarly, the number of hits according to selected algorithm vs. number of apples from visual counting. Such analysis allows us to determine the coefficient of determination ( $R^2$ ). As data analysis, errors of omission, commission, and detection accuracy were calculated for each image using equations 11 to 13 (adapted from Congalton, 1991).

$$Omission = \frac{FN}{TP} \quad (11)$$

$$Commission = \frac{FP}{TP} \quad (12)$$

$$Accuracy = \frac{TP}{TP + FN} \quad (13)$$

where: FN – false negative; TP – true positive and FP – false positive

Omission error values are those that do not differentiate the apple fruit from the background environment, failing to mark them. On the other hand, commission errors occur when detecting features in the environment that do not match the fruits, and these are accounted for.

## 3. Results

The images acquired in the field resulted in sampling seven entire plants. Considering the overlapping of the scenes, a total of 20 images were required, ten from each side of the selected planting row (Figure 3). The average time for the image processing for all space colors till the fruit counting is a process that takes 12.4 seconds. Color transformation (for example *Lab-a*) till the fruit counting takes only 4.2 seconds per image, what resulted in total time of 84 seconds. As a result, a total of 4,056 visible fruits were counted (Table 2).

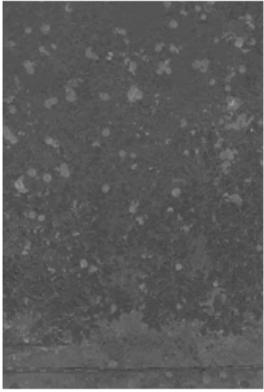
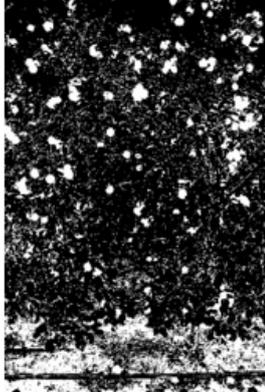
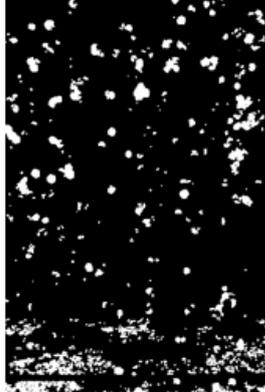
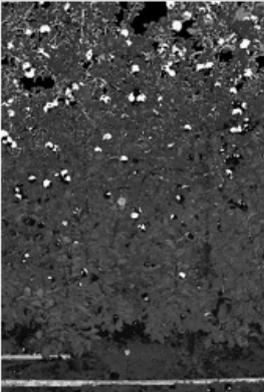
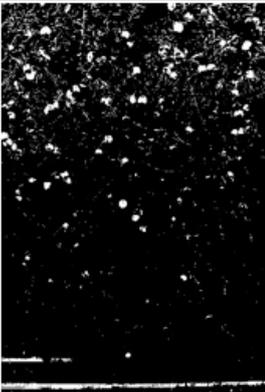
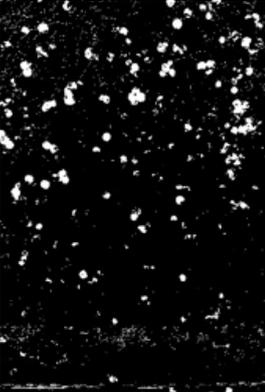
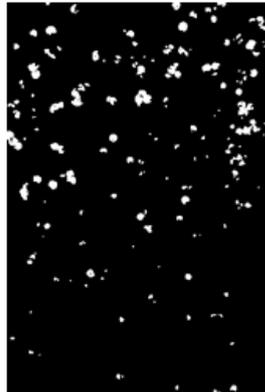
**Table 2:** Sample values after image acquisition and visual fruit count.

Samples	Total Number
Images	20
Apple tree	7
Apple fruit VC*	4,056

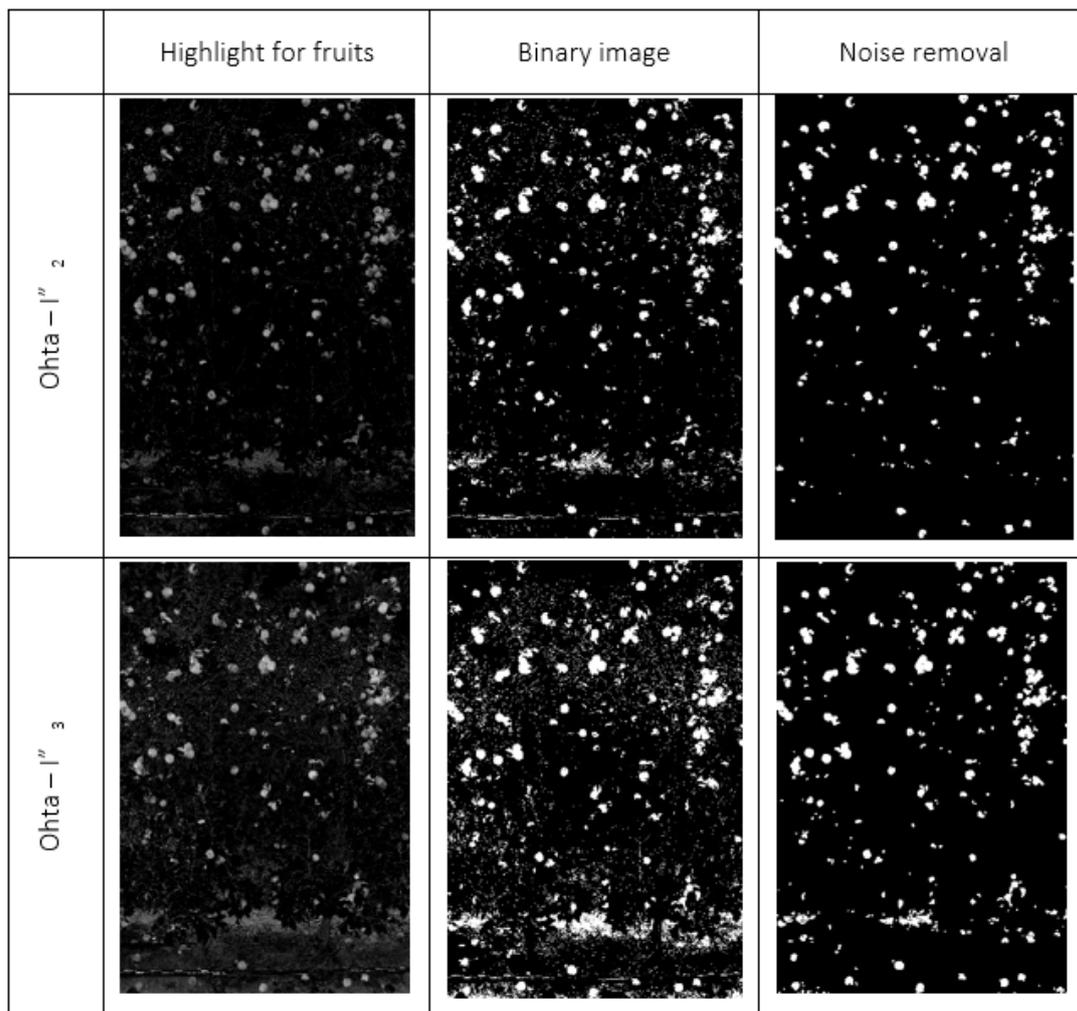
\* VC = Visual Count

We use the subset of the terrestrial RGB image shown in Figure 1E as an example to highlight the different digital image processing steps proposed in this study. Figures 6 and 7 show both the segmentation and noise removal stage, using the color space transformation methodology. Only the bands that showed enhancement for the fruits

were selected, namely  $r$  from normalized  $rgb$  ( $rgb-r$ ),  $h$  from HSV space ( $HSV-h$ ),  $a$  from  $L^*a^*b^*$  space ( $Lab-a$ ), and bands  $I''_2$  and  $I''_3$  of Ohta space (Figure 6 and Figure 7).

	Highlight for fruits	Binary image	Noise removal
$rgb-r$			
$HSV-h$			
$L^*a^*b^*-a$			

**Figure 6:** Chart of the results of color space transformations, image binarization and noise removal. Displayed bands:  $rgb-r$ ,  $HSV-h$  and  $L^*a^*b^*-a$ .



**Figure 7:** Chart of the results of Ohta color space transformations, image binarization and noise removal. Displayed bands:  $I''_2$  and  $I''_3$ .

The analysis of the images generated after the filtering process shows that some bands ( $rgb-r$ ,  $HSV-h$  and  $I''_3$ ) continue to have noise at the bottom of the image, corresponding to the ground and the reference scale stadia. A simple solution to this problem was making combinations between bands using binary images. The combinations performed were:

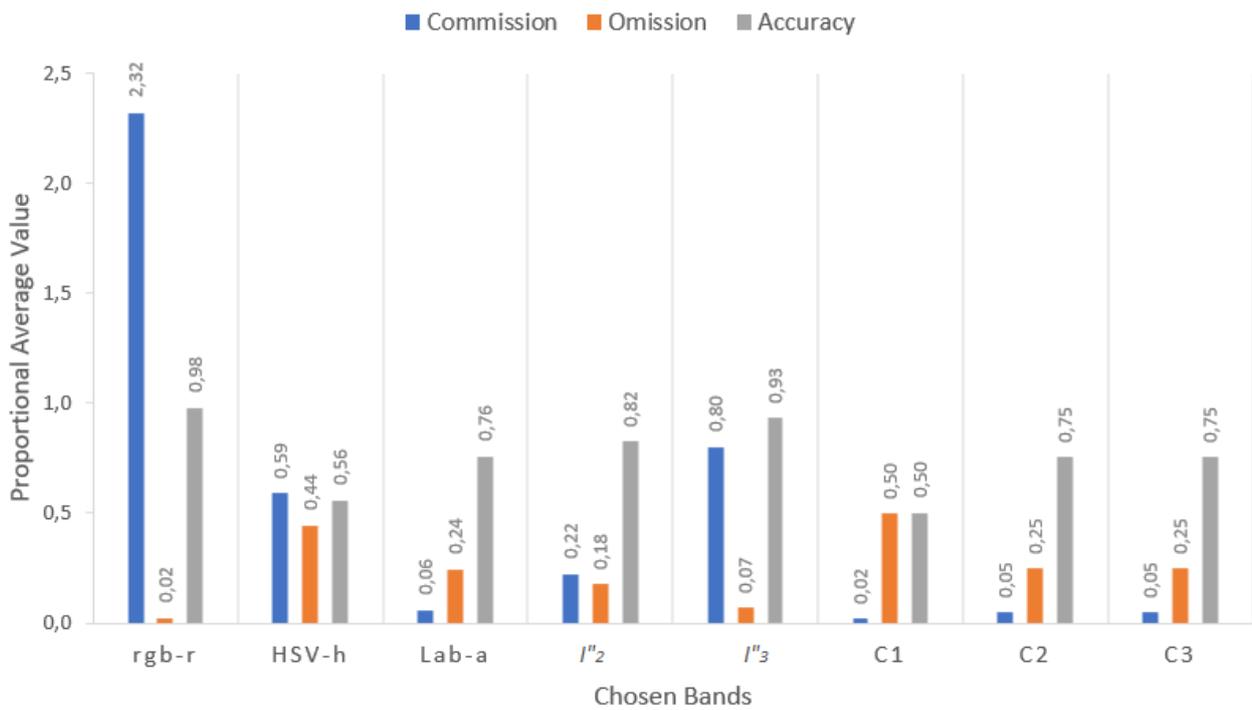
$$\begin{cases} C1 = h + a + I''_3 \\ C2 = r + a + I''_3 \\ C3 = a + I''_2 + I''_3 \end{cases} \quad (14)$$

The new three images generated were added to the erosion and dilation process. The selected bands and the combined images proceeded to the counting step of the structures identified as fruits. The mean values of the 20 images are presented in Table 3.

**Table 3:** Mean values of visual fruit counting, automatic counting, commission, omission, and accuracy error values for 20 images and for images from separate sides, W and E.

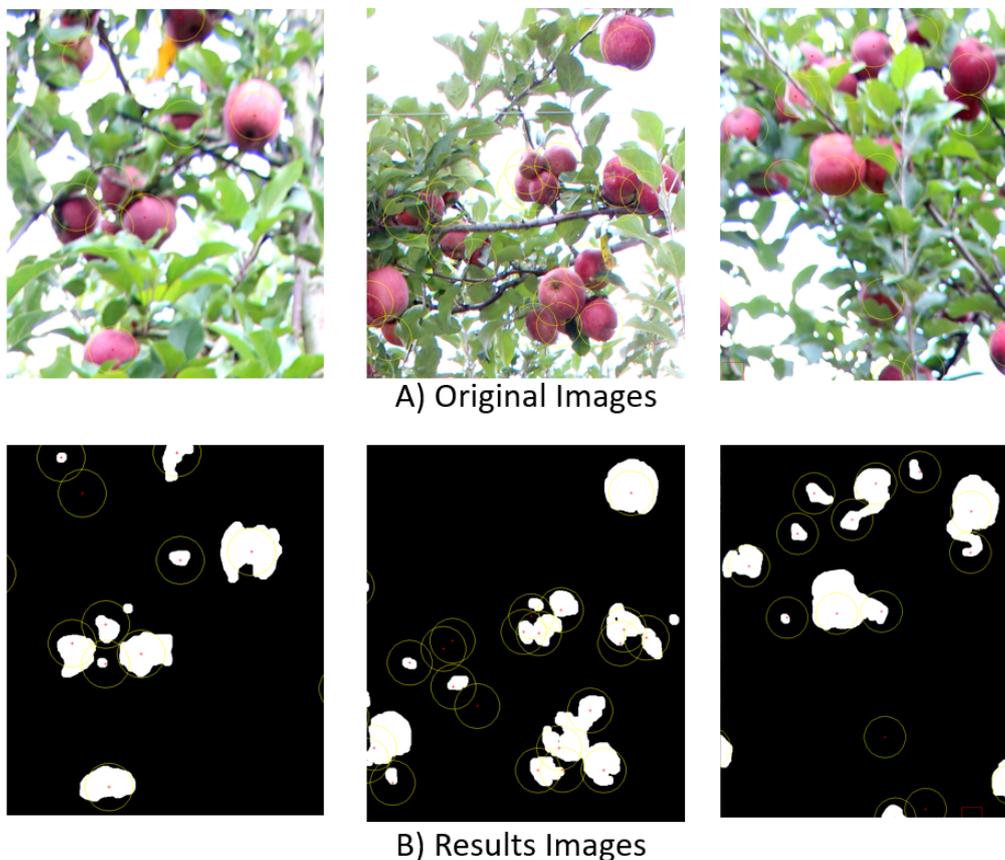
		Bands								
N = 20 images		RGB	rgb-r	HSV-h	Lab-a	I'' <sub>2</sub>	I'' <sub>3</sub>	C1	C2	C3
VC	μ	203	-	-	-	-	-	-	-	-
	σ	45.5	-	-	-	-	-	-	-	-
AC	μ	-	632	210	144	185	318	93	142	142
	σ	-	129.5	58.0	23.0	27.0	32.0	13.7	24.3	23.7
Commission	μ	-	2.32	0.59	0.06	0.22	0.80	0.02	0.05	0.05
	σ	-	1.18	0.42	0.03	0.18	0.31	0.01	0.02	0.02
Omission	μ	-	0.02	0.44	0.24	0.18	0.07	0.50	0.25	0.25
	σ	-	0.02	0.07	0.03	0.03	0.02	0.07	0.03	0.03
Accuracy	μ	-	0.98	0.56	0.76	0.82	0.93	0.50	0.75	0.75
	σ	-	0.02	0.07	0.03	0.03	0.02	0.07	0.03	0.03
		Side W								
N = 10 images		RGB	rgb-r	HSV-h	Lab-a	I'' <sub>2</sub>	I'' <sub>3</sub>	C1	C2	C3
VC	μ	243 a*	-	-	-	-	-	-	-	-
	σ	24.9	-	-	-	-	-	-	-	-
AC	μ	-	575	191	163	193	323	102	163	162
	σ	-	108.6	58.5	16.2	24.1	36.9	6.8	17.1	16.4
Commission	μ	-	1.53	0.33	0.04	0.11	0.55	0.01	0.04	0.04
	σ	-	0.53	0.22	0.01	0.07	0.18	0.01	0.01	0.01
Omission	μ	-	0.04	0.48	0.24	0.17	0.07	0.53	0.25	0.25
	σ	-	0.02	0.06	0.04	0.04	0.03	0.05	0.04	0.04
Accuracy	μ	-	0.96	0.52	0.76	0.83	0.93	0.47	0.75	0.75
	σ	-	0.02	0.06	0.04	0.04	0.03	0.05	0.04	0.04
		Side E								
N = 10 images		RGB	rgb-r	HSV-h	Lab-a	I'' <sub>2</sub>	I'' <sub>3</sub>	C1	C2	C3
VC	μ	162 b*	-	-	-	-	-	-	-	-
	σ	13.6	-	-	-	-	-	-	-	-
AC	μ	-	689	229	124	176	312	83	121	121
	σ	-	123.5	50.6	5.2	27	25	12.2	6	6
Commission	μ	-	3.11	0.85	0.07	0.34	1.05	0.03	0.06	0.06
	σ	-	1.12	0.41	0.03	0.18	0.17	0.02	0.03	0.03
Omission	μ	-	0.01	0.40	0.24	0.18	0.07	0.47	0.24	0.24
	σ	-	0.01	0.05	0.03	0.03	0.01	0.07	0.02	0.02
Accuracy	μ	-	0.99	0.60	0.76	0.82	0.93	0.53	0.76	0.76
	σ	-	0.01	0.05	0.03	0.03	0.01	0.07	0.02	0.02

VC=Visual Count; AC=Automatic Counting; \* Means followed by the same letter do not differ significantly from each other at 5% probability by Tukey's test.



**Figure 8:** Comparison graph of the general mean of commission, omission and accuracy images by detection method.

An example of the results of detections is shown in Figure 9. With visualization of success and failure in detections.



**Figure 9:** Visual fruit counting in original images (A) and results of detections with the *Lab-a* image algorithm (B).

## 4. Discussion

Bands that presented commission errors with values higher than 20% (0.2) were: *rgb-r*, *HSV-h*,  $I''_2$  and  $I''_3$ . These high error values exclude the possibility of using a single band for apple fruit detection. The commission error brings more problems because false positives are being computed for apple fruits. It is observed that *rgb-r* showed higher objects detections than the number of fruits exist in the images. Most of this error occurred due to the fraction of exposed soil presented in the interline that appears in the images; the same problem was detected with the use of the band  $I''_3$ .

The bands with omission errors corresponding to values greater than 20% (0.2) were: *HSV-h*, *Lab-a*, *C1*, *C2*, and *C3*. In this case, the dynamic threshold adopted failed to differentiate some fruits from the complex background environment. Wei et al. (2014) was also unable to identify unripe fruits and ripe fruits. This behavior was also due to a high occlusion rate that hindered the proper luminosity of the fruit. This type of error was greater in fruits that were in the lower portion of the plant and inside the inner layer of the plant.

The degree of accuracy of hits in the bands is also presented. Hits in relation to the total of apple fruits in the image that exceeded 75% (0.75) occurred with: *rgb-r*, *Lab-a*,  $I''_2$ ,  $I''_3$ , *C2* and *C3*. The highest hit values were *rgb-r*,  $I''_2$  e  $I''_3$ , all above 80% (0.8). However, these bands also presented high values of commission error. In practice, this makes the use of these single-bands unfeasible as an autonomous method of detecting apple fruits under the conditions tested in this study. Figures 6 and 7 show in detail the detection of the apple fruits in these bands. However, the greatest source of error occurred in fruits located over the ground. Further studies shall consider changing the angle of the sensor so that the field of view does not detect the ground layer reducing therefore the commission values in these bands.

Thus, the only bands that remained in the analysis were *Lab-a*, *C2* and *C3*. The three presented practically identical values of commission (0.06; 0.05; 0.05), omission (0.24; 0.25; 0.25) and accuracy (0.76; 0.75; 0.75), respectively. These findings corroborate the study performed by Linker and Kelman (2015). These authors obtained 0.76  $R^2$  value in apple detection, with a 95% hit rate when compared to the fruits counted in the image. Their commission, omission and accuracy values were 16%, 28% and 72%, respectively. Values very similar to those found for *Lab-a* band in this study.

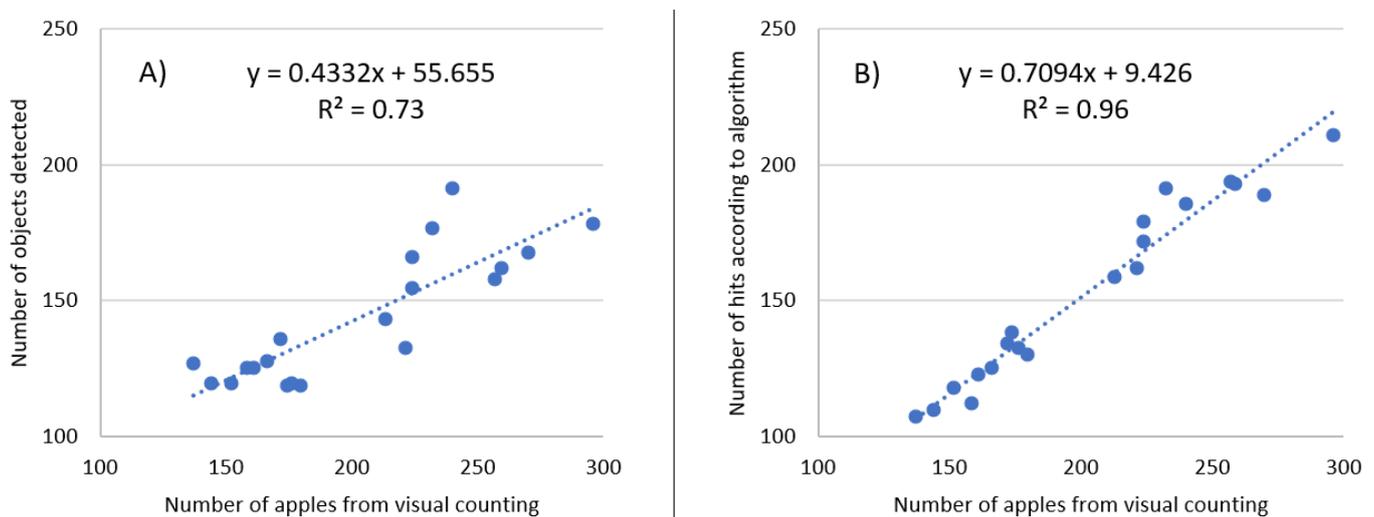
Gongal et al. (2016) described 78.9% in accuracy of hits in detecting apple fruits in full plant evaluation using a set of 2D and 3D cameras. However, the slight improvement implies a higher cost design and more complex data processing procedures. Zhou et al. (2012) present a 0.85 correlation coefficient between manual fruit counting and counting by the color characteristic algorithm in the period before harvest.

Studies using more sophisticated detection techniques, such as deep learning, have shown more accurate results in the identification of fruits with close range images. However, deep learning requires a higher time of annotation and training, as well as high computational processing costs by training a large dataset (Häni et al. 2020, Apolo-Apolo et al. 2020). Although suggested for future studies, the financial cost of advanced computers is probably unaffordable for most small holder farmers when we consider a short timeframe. Additionally, it would be only feasible if datasets are labeled are provided for training and validating. On the other hand, the main advantage of this method rely on the use on original RGB images.

Other studies that do not consider the whole plant, with the sole purpose of identifying the fruit for robotic vision, and analyzing the fruit set in isolation, achieved higher hit rates such as 86-100% (Jidong et al., 2016) and 95-100% (Wei et al., 2014). The combined bands (*C2* and *C3*) derive from a combination with *Lab-a*, which proves to be the limiting variable in these combinations. Therefore, because it originates from a direct transformation and does not present differences of the combined bands (*C2* and *C3*), the use of *Lab-a* band is an alternative for the detection of Fuji Suprema apple fruits and limited to the same conditions of this study.

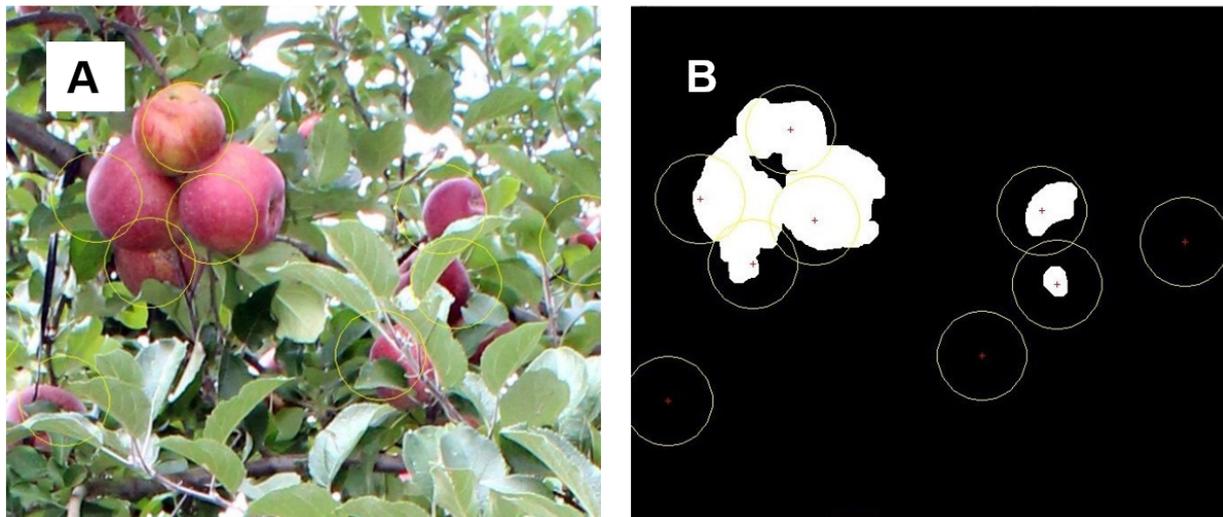
The component  $a$  of the  $Lab$  represents values on an axis that vary between red (positive values) and green (negative values), and in component  $b$  the values vary between yellow (positive values) and blue (negative values). As the apple fruits strongly emphasize the red color and both the leaves and ground coverage tend to green, the differentiation on the same axis facilitates the division of these objects in the image (Kahu, S.Y., Raut, R.B., and Bhurchandi, K.M., 2019).

Applying a regression analysis between the number of detected objects by connected components and the number of fruits of the visual count resulted in a coefficient of determination ( $R^2$ ) of 0.73, but when compared with the crossing between the number of hits of the algorithm and the number of fruits of the visual count,  $R^2$  is 0.96 (Figure 10). This was because the algorithm could not segment fruit clusters into individual fruits, and the resulted counting of the detected objects was through the method of connected components. Considering this cluster as just a detected object, where there were two or more fruits (Figure 11 A and B). To improve hit rates in fruit detection in images, future projects should focus on fruit individualization and separation, using techniques that, in addition to exploring color spaces, consider other attributes of the fruit such as shape and texture. Studies on other orchard farms and under different lighting conditions are also recommended to confirm the results achieved in this study. Finally, sensor systems idealized for harvesting purposes should be calibrated with other agronomic parameters so that more reliable estimations can be provided.



**Figure 10:** Regressions between number of fruits counted visually in the RGB image and: the number of objects detected in the algorithm in the  $Lab-a$  image (A); the number of hits of the algorithm (B).

We also intend to optimize the entire process in near future. Improvements are expected so that it can be implemented in PC and in smartphones. The advantage is that the counting process can be performed easily using such devices and providing results in real time and at low cost. The disadvantages are still the need for the quality inspection from the user that still has to judge the best bands for the fruit detection since the proposed method is based purely on color spaces.



**Figure 11:** Fruits clustered (A) and counted as a single object by having the edges together (B).

## 5. Conclusion and Recommendations

The detection of ripe apple fruits before harvesting by our semi-automatic algorithm using low-cost RGB images performed similarly to those approaches reported in the literature using a background screen.

Failure to use a background screen could pose major problems of detection confusion, but *Lab-a* color space transformation has addressed this adversity. The detection rate ( $R^2=0.73$ ) was compatible with other studies, as well as the fruit recognition accuracy rate ( $R^2=0.96$ ). Experiments employing other environmental conditions and orchards are recommended for future research projects.

Unripe fruits with poor lighting were not detected in the proposed methodology, which is a use limitation in this case. Techniques such as deep learning that rely on original RGB images can be used to solve the problem and are also suggested for future projects.

The authors believe that the accuracy indexes may improve under plant conduction systems that leave the fruits more exposed. This may contribute to the reduction of the tree canopy volume, giving, therefore, a characteristic of the two-dimensional form. This may contribute to an increase in the fruit accuracy rate by decreasing the omission error.

High early detection values relative to harvest improve yield prediction rates. This study thus contributed to a method of low user intervention and simple processing to detect ripe apple fruits. Improvements should be made to differentiate fruit clusters and detect unripe fruits individually.

Production estimate calculations should also be explored in future studies to assist in decision making by orchard managers, as well as exploring the use of color segmentation in fructiculture, whether in detecting fruits, leaves, trunks and twigs or injury caused by disease and pests. Tests with other color spaces combined with deep learning techniques are promising alternatives for fruit research. Finally, improvements in presenting the proposed approach inside user-friendly tools allowing their practical use in either PC or smartphones shall be also explored in near future.

## ACKNOWLEDGMENTS

The authors would like to thank the State University of Santa Catarina (UDESC), and the department of

Environmental and Sanitation Engineering, for supporting the doctoral dissertation of the first author. We would also like to thank the owner of the study site for allowing us access into the rural property, as well as for their availability and generosity. This research was partially funded by the Santa Catarina Research Foundation (FAPESC; 2017TR1762, 2019TR816), the Brazilian National Council for Scientific and Technological Development (CNPq; 307689/2013-1, 303279/2018-4, 313887/2018-7 and 436863/2018-9)

## AUTHOR'S CONTRIBUTION

Conceptualization: L.J.B; Methodology: L.J.B., V.L., J.A.S.C.; Data collection: L.J.B., V.L., M.B.S.; Implementation and data processing: L.J.B., V.L., J.A.S.C.; Writing – original draft: L.J.B.; Supervision: E.A.M., J.A.S.C., V.L.; Discussion of the results, paper revision and end edition: L.J.B., E.A.M., V.L., L.R.

## REFERENCES

- An, N. *et al.* 2017. Quantifying time-series of leaf morphology using 2D and 3D photogrammetry methods for high-throughput plant phenotyping, *Computers and Electronics in Agriculture*, 135, pp. 222–232. doi: 10.1016/j.compag.2017.02.001.
- Apolo-Apolo, O. E. *et al.* 2020., A Cloud-Based Environment for Generating Yield Estimation Maps From Apple Orchards Using UAV Imagery and a Deep Learning Technique, *Frontiers in Plant Science*, 11(7), pp. 1–15. doi: 10.3389/fpls.2020.01086.
- Aquino, A. *et al.* 2018. Automated early yield prediction in vineyards from on-the-go image acquisition, *Computers and Electronics in Agriculture*, 144, pp. 26–36. doi: 10.1016/j.compag.2017.11.026.
- Bargoti, S. and Underwood, J. P. 2017. Image Segmentation for Fruit Detection and Yield Estimation in Apple Orchards, *Journal of Field Robotics*, 34(6), pp. 1039–1060. doi: 10.1002/rob.21699.
- Behroozi-Khazaei, N. and Maleki, M. R. 2017. A robust algorithm based on color features for grape cluster segmentation, *Computers and Electronics in Agriculture*, 142, pp. 41–49. doi: 10.1016/j.compag.2017.08.025.
- Berk, P. *et al.* 2016. Development of alternative plant protection product application techniques in orchards, based on measurement sensing systems: A review, *Computers and Electronics in Agriculture*, 124, pp. 273–288. doi: 10.1016/j.compag.2016.04.018.
- Bittencourt, C. C. *et al.* 2011. A cadeia produtiva da maçã em Santa Catarina: competitividade segundo produção e packing house, *Revista de Administração Pública*, 45(4), pp. 1199–1222.
- Cheng, H. *et al.* 2017. Early Yield Prediction Using Image Analysis of Apple Fruit and Tree Canopy Features with Neural Networks, *Journal of Imaging*, 3(1), p. 6. doi: 10.3390/jimaging3010006.
- Coelho Filho, M. A. *et al.* 2005. Estimativa da área foliar de plantas de lima ácida 'Tahiti' usando métodos não-destrutivos, *Revista Brasileira de Fruticultura*, 27(1), pp. 163–167. doi: 10.1590/s0100-29452005000100043.
- Colaço, A. F. *et al.* 2017. A method to obtain orange crop geometry information using a mobile terrestrial laser scanner and 3D modeling, *Remote Sensing*, 9(8), pp. 10–13. doi: 10.3390/rs9080763.
- Congalton, R. G. 1991. A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sensing of Environment*, 37(1), pp. 35–46. doi: 10.1016/0034-4257(91)90048-B
- Dorj, U. O., Lee, M. and Yun, S. 2017. An yield estimation in citrus orchards via fruit detection and counting using image processing, *Computers and Electronics in Agriculture*, 140, pp. 103–112. doi: 10.1016/j.compag.2017.05.019.

- Escolà, A. et al. 2017. Mobile terrestrial laser scanner applications in precision fruticulture/horticulture and tools to extract information from canopy point clouds, *Precision Agriculture*, 18(1), pp. 111–132. doi: 10.1007/s11119-016-9474-5.
- Font, D. et al. 2014a. Counting red grapes in vineyards by detecting specular spherical reflection peaks in RGB images obtained at night with artificial illumination, *Computers and Electronics in Agriculture*, 108, pp. 105–111. doi: 10.1016/j.compag.2014.07.006.
- Font, D. et al. 2014b. A proposal for automatic fruit harvesting by combining a low cost stereovision camera and a robotic arm, *Sensors*, 14(7), pp. 11557–11579. doi: 10.3390/s140711557.
- Gené-Mola, J. et al. 2019. Multi-modal deep learning for Fuji apple detection using RGB-D cameras and their radiometric capabilities, *Computers and Electronics in Agriculture*, 162, pp. 689–698. doi: 10.1016/j.compag.2019.05.016.
- Gongal, A. et al. 2016. Apple crop-load estimation with over-the-row machine vision system, *Computers and Electronics in Agriculture*, 120, pp. 26–35. doi: 10.1016/j.compag.2015.10.022.
- Häni, N., Roy, P. and Isler, V. 2020. A comparative study of fruit detection and counting methods for yield mapping in apple orchards, *Journal of Field Robotics*, 37(2), pp. 263–282. doi: 10.1002/rob.21902
- IBGE, 2018. *Produção Agrícola Municipal 2017*, IBGE. Available at: <<https://cidades.ibge.gov.br/brasil/sc/pesquisa/15/11863?ano=2017>> [Accessed: September 28, 2019].
- Jidong, L. et al. 2016. Recognition of apple fruit in natural environment, *Optik*, 127(3), pp. 1354–1362. doi: 10.1016/j.ijleo.2015.10.177.
- Kahu, S.Y., Raut, R.B., and Bhurchandi, K.M. 2019. Review and evaluation of color spaces for image/video compression, *Color Research and Application*, 44(1), pp. 8–33.
- Kang, H. and Chen, C. 2019. Fruit Detection and Segmentation for Apple Harvesting Using Visual Sensor in Orchards, *Sensors*, 19(20), p. 4599. doi: 10.3390/s19204599.
- Koenig, K. et al. 2015. Comparative classification analysis of post-harvest growth detection from terrestrial LiDAR point clouds in precision agriculture, *ISPRS Journal of Photogrammetry and Remote Sensing*. International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS), 104, pp. 112–125. doi: 10.1016/j.isprsjprs.2015.03.003.
- Koirala, A. et al. 2019. Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of ‘MangoYOLO,’ *Precision Agriculture*, 20(6), pp. 1107–1135. doi: 10.1007/s11119-019-09642-0.
- Liang, X. et al. 2014. The use of a hand-held camera for individual tree 3D mapping in forest sample plots, *Remote Sensing*, 6(7), pp. 6587–6603. doi: 10.3390/rs6076587.
- Linker, R. 2017. A procedure for estimating the number of green mature apples in night-time orchard images using light distribution and its application to yield estimation, *Precision Agriculture*, 18(1), pp. 59–75. doi: 10.1007/s11119-016-9467-4.
- Linker, R. and Kelman, E. 2015. Apple detection in nighttime tree images using the geometry of light patches around highlights, *Computers and Electronics in Agriculture*, 114, pp. 154–162. doi: 10.1016/j.compag.2015.04.005.
- Liu, X. et al. 2016. A method of segmenting apples at night based on color and position information, *Computers and Electronics in Agriculture*, 122, pp. 118–123. doi: 10.1016/j.compag.2016.01.023.
- Liu, X. et al. 2019. A Detection Method for Apple Fruits Based on Color and Shape Features, *IEEE Access*, 7, pp. 67923–67933. doi: 10.1109/ACCESS.2019.2918313.
- Ohta, Y.-I., Kanade, T. and Sakai, T. 1980. Color information for region segmentation, *Computer Graphics and Image Processing*. Academic Press, 13(3), pp. 222–241. doi: 10.1016/0146-664X(80)90047-7.
- Otsu, N. 1979. A Threshold Selection Method from Gray-Level Histograms, *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), pp. 62–66. doi: 10.1109/TSMC.1979.4310076.

Petri, J. L. *et al.* 2011. AVANÇOS NA CULTURA DA MACIEIRA NO BRASIL, *Revista Brasileira de Fruticultura*. Especial, pp. 048–056.

Shaik, K. B. *et al.* 2015. Comparative Study of Skin Color Detection and Segmentation in HSV and YCbCr Color Space, *Procedia Computer Science*, 57, pp. 41–48. doi: 10.1016/j.procs.2015.07.362.

Tao, Y. and Zhou, J. 2017. Automatic apple recognition based on the fusion of color and 3D feature for robotic fruit picking, *Computers and Electronics in Agriculture*, 142, pp. 388–396. doi: 10.1016/j.compag.2017.09.019.

Vázquez-Arellano, M. *et al.* 2016. 3-D imaging systems for agricultural applications—a review, *Sensors*, 16(5). doi: 10.3390/s16050618.

Wei, X. *et al.* 2014. Automatic method of fruit object extraction under complex agricultural background for vision system of fruit picking robot, *Optik*, 125(19), pp. 5684–5689. doi: 10.1016/j.ijleo.2014.07.001.

Zhou, R. *et al.* 2012. Using colour features of cv. ‘Gala’ apple fruits in an orchard in image processing to predict yield, *Precision Agriculture*, 13(5), pp. 568–580. doi: 10.1007/s11119-012-9269-2.