IMPROVEMENT PROCEDURE FOR IMAGE SEGMENTATION OF FRUITS AND VEGETABLES BASED ON THE OTSU METHOD

OSBALDO VITE–CHÁVEZ[1], JORGE FLORES–TRONCOSO[1], REYNEL OLIVERA–REYNA[2] AND JORGE ULISES MUÑOZ-MINJARES[✉,2]

[1]Departamento de posgrado en Ingeniería y Tecnología Aplicada, Unidad Académica de Ingeniería Eléctrica, Universidad Autónoma de Zacatecas, Av. Ramón López Velarde, 801, Col. Centro, 98000, Zacatecas, México, [2]Departamento de Comunicaciones y Electrónica, Unidad Académica de Ingeniería Eléctrica, Universidad Autónoma de Zacatecas, Campus Jalpa, Lib. Jalpa Km. 156+380, Fracc. Solidaridad, Carr. Gdl.-Sal., Jalpa, 99601, Zac., México.
e-mail: osvichz@uaz.edu.mx, jflorest@uaz.edu.mx, reynel@uaz.edu.mx, ju.munoz@uaz.edu.mx.

ABSTRACT

Currently, there are significant challenges in the classification, recognition, and detection of fruits and vegetables. An important step in solving this problem is to obtain an accurate segmentation of the object of interest. However, the background and object separation in a grayscale image shows high errors for some thresholding techniques due to uneven or poorly conditioned lighting. An accepted strategy to reduce segmentation errors is to select the channel of an RGB image with high contrast. This paper presents the results of an experimental procedure based on enhancing binary segmentation by using the Otsu method. The procedure was carried out with images of real agricultural products, both with and without additional noise, to corroborate the robustness of the proposed strategy. The experimental tests were performed using our database of RGB images of agricultural products under uncontrolled illumination. The results show that the best segmentation is achieved by selecting the Blue channel of the RGB test images due to its higher contrast. Here, the quantitative results are measured by applying the Jaccard and Dice metrics based on the ground-truth images as optimal reference. Most of the results using both metrics show an improvement greater than 45.5% in the two experimental tests.

Keywords: contrast, fruits and vegetable images, illumination, improvement, Otsu, segmentation.

INTRODUCTION

In marketing agricultural products, the quality evaluation is realized according to their properties such as color, shape, and size. These products are visually inspected and are time-consuming, thus generating high operational costs. This strategy is susceptible to human error, and it could present difficulties in standardizing the classification results. The selection error of an agricultural product can affect its quality and commercial value (Yuan *et al.*, 2015; Belan *et al.*, 2020; Mukhiddinov *et al.*, 2022). A frequent problem identified in vision systems and intelligent applications, in processing images of agricultural products, is the uncontrolled illumination of the environment. This issue generates images with low contrast, saturated color areas, and problems with brightness and shadows. Consequently, these factors hide information for any inspection algorithm to provide a proper solution in product classification (Yuan *et al.*, 2015; Alegre *et al.*, 2016). Dim illumination can lead to a low signal-to-noise ratio, which can affect an image with noisy pixels (Gonzalez *et al.*, 2008; Russ, 2016). Also, uneven lighting can significantly hinder segmentation operations when the thresholding techniques, such as Otsu's method, are applied (Ng, 2006; Gonzalez *et al.*, 2008; Yuan *et al.*, 2015; Alegre *et al.*, 2016).

Image segmentation is a crucial step for analyzing and interpreting acquired image in various fields such as medical imaging, agriculture, robotic vision, material science, geographical imaging, and more. (Sha *et al.*, 2016; Lu *et al.*, 2017; Goh *et al.*, 2018; Resma *et al.*, 2018; Song *et al.*, 2020). To enhance the efficiency of a vision system, adequate lighting or an optimal segmentation algorithm is required to obtain the correct image for further processing. Consequently, a high-quality image can provide the necessary information for detecting and identifying objects, along with their relevant characteristics(Alegre *et al.*, 2016). An ideal image for object segmentation is one in which the pixels representing the object of interest share similar brightness characteristics, distinct from the background (Gonzalez *et al.*, 2008; Russ, 2016). Nowadays, numerous segmentation methods have been proposed due to the nonideal nature of real images caused by a multitude of variables.

Thresholding is a popular segmentation technique for several applications due to its simplicity and efficiency (Liu *et al.*, 2015; Sha *et al.*, 2016; Goh

et al., 2018; Lei *et al.*, 2019; Song *et al.*, 2020). Basically, two types of thresholding processes exist: *global thresholding* and *local thresholding*. Global thresholding involves selecting a single value from the histogram of the entire image. On the other hand, local thresholding uses information from the gray intensity level of the histogram to determine several thresholding values. The implementation of global thresholding is simpler and easier, but its result rely on good illumination, i.e. uniform illumination (Ng, 2006; Yuan *et al.*, 2015). Local thresholding methods yield the best results in images of texts or manuscripts with non-uniform illumination, but their algorithms require more computational processing, leading to slower processing (Gatos *et al.*, 2006; Singh *et al.*, 2012).

The Niblack (1986) and Sauvola *et al.* (2000) binarization algorithms are among the most sophisticated techniques that employ local thresholds. These proposed algorithms obtain several local thresholds for the neighborhoods or subregions of the image using a sliding window to calculate variables such as the mean, variance, and standard deviation. Gatos *et al.* (2006) proposed an adaptable methodology for the binarizing documents degraded by shadows, non-uniform illumination, low contrast, and other factors. In their proposal, they apply a pre-processing using a Weiner filter to estimate the background and apply thresholding using the Sauvola technique. Bradley *et al.* (2007) proposed an adaptive threshold search algorithm by calculating the mean in each local neighborhood.

In this approach, the value of the local threshold depends on local statistics, such as range, variance, or pixel neighborhood surface fit parameters. However, while these methods are tolerant to illumination changes, they can be sensitive to noise, potentially degrading the final segmentation output (Bataineh *et al.*, 2017; Cheremkhin *et al.*, 2019). Global thresholding is commonly used for its simplicity and speed in various applications, such as automated visual inspection, where the illumination conditions are controlled, i.e., uniform illumination (Sezgin *et al.*, 2004; Ng, 2006; Liu *et al.*, 2015).

Among the global thresholding techniques, Sahoo *et al.* (1988) concluded that the method described by Otsu Otsu (1979) is one of the best threshold selection methods for real-world images concerning uniformity and shape measures. This method selects one or several threshold values that maximize the inter-class variations of the histogram. However, it is only optimal for setting thresholds of a histogram with bimodal or multimodal distribution, i.e., for images where object and background distributions are clearly defined in shape and size. Therefore, the Otsu method does not yield the desired results if the histogram has a unimodal or near-unimodal distribution (Liu *et al.*, 2015; Lei *et al.*, 2019). In Chávez *et al.* (2022), it was verified that thresholding algorithms obtain better results when the images show higher contrast between background and object.

Commonly, thresholding algorithms are used for product quality inspection using computer vision (Fan *et al.*, 2021). The basic idea is to automatically obtain an optimal gray-level threshold value to separate the object of interest from the background. Thus, the threshold value can be obtained based on the gray level distribution in the histogram of a digital image (Goh *et al.*, 2018; Resma *et al.*, 2018; Lei *et al.*, 2019). Aiming to improve the results of binary segmentation in fruits and vegetable images by applying the Otsu method under uncontrolled lighting environment, the following procedure is proposed under the premise that gray images with higher contrast yield better results in a binary segmentation process. The manuscript is structured as follows to corroborate this hypothesis. First, Section 2 describes the basic concepts, methods, and metrics. Next, two experimental tests with and without additional noise, are carried out, and two operations are applied by using the luminance equation and selecting the blue channel from all RGB images in Section 3. Results and Discussion are detailed in Section 4. Finally, Section 5 presents the conclusions.

THEORETICAL FOUNDATIONS

In this section, the principal concepts and techniques employed in this article are described in detail. The nomenclature used throughout the document was described in the framed nomenclature in this section.

DIGITAL IMAGE

An image is defined as a function $f(x,y)$ in two-dimension (*2D*), where x and y are spatial coordinates in the (x,y) plane. The function $f(x,y)$ at any point is called the intensity of the image with a range of $[0, L-1]$, where $L = 256$. When the intensity values of $f(x,y)$ are finite and discrete quantities, it can be defined as a digital image. This is composed of a finite number of elements called pixels, with a particular location (x,y) and a corresponding value (Gonzalez *et al.*, 2008; Sundararajan, 2017).

Meanwhile, an RGB image can be composed with three main channels $I_{RGB} = [R, G, B]$, where R, G, and B are the pixel intensity in the Red, Green, and Blue images, respectively. Therefore, an image in the RGB color space is represented by three image

	Nomenclature
I_{RGB}	RGB image.
$I_{g(x,y)}$	Gray scale image
$\dot{I}_g$	Gray image from equation (1).
$\ddot{I}_g$	Gray image from channel selection.
$\dot{I}_{gn}$	Gray image with noise from equation (1).
$\ddot{I}_{gn}$	Gray image with noise from channel selection.
$\dot{I}_G$	Binary image from equation (1) with matlab command `graythresh`.
$\dot{I}_O$	Binary image from equation (1) with matlab command `otsuthresh`.
$\ddot{I}_G$	Binary image from channel selection with matlab command `graythresh`.
$\ddot{I}_O$	Binary image from channel selection with matlab command `otsuthresh`.
$\dot{I}_{Gn}$	Binary image from eq. (1) given a noise vector with matlab command `graythresh`.
$\dot{I}_{On}$	Binary image from eq. (1) given a noise vector with matlab command `otsuthresh`.
$\ddot{I}_{Gn}$	Binary image from channel selection given a noise vector with matlab command `graythresh`.
$\ddot{I}_{On}$	Binary image from channel selection given a noise vector with matlab command `otsuthresh`.

components of size $M \times N$ with conventional brightness intensities between 0 and 255. Frequently, the term grayscale is employed in digital images, which means that each pixel value represents only the light intensity information. These grayscale images can be defined as $I_{g(x,y)} \in \mathbb{Z}_G^{M \times N}$, this function shows values from the darkest black to the brightest white, going through various levels of gray intensity. Here, G represents the maximum base value 2^k and $k = 8 \ bits$. That is, the gray value of the image is usually represented by the combination of eight binary numbers that constitute the value of a pixel (Gonzalez *et al.*, 2008; Sundararajan, 2017).

An RGB image is converted to grayscale by adding the levels of each channel in different percentages of Red (30%), Green (59%), and Blue (11%). This procedure is based on how the human eye perceives the frequencies of the spectrum nearby to the light intensity of these primary colors (Bovik, 2009; Russ, 2016). This phenomenon and the weighting factors of each color component can be expressed in the mathematical equation of luminance,

$$GS = 0.2989 * R + 0.5870 * G + 0.1140 * B. \quad (1)$$

Therefore, the equation (1) is applied to each pixel in the RGB image to convert an image from RGB to grayscale. The result is a new matrix $\dot{I}_{g(x,y)}$ of one byte per pixel that provides the luminance information. This weighting is a single intensity shade of gray (Bovik, 2009; Russ, 2016).

Now, a binary image is a digital image that can be obtained based on the matrix $\dot{I}_{g(x,y)}$. The binary image $I_{b(x,y)}$ is a digital image that has two possible values for each pixel $[0, 1]$, of size $M \times N$, such that $I_b \in \mathbb{Z}_2^{M \times N}$. Normally, the colors used for the representation of a binary image are black and white, although any pair of colors can be used. So, one of the colors represents the background and the other represents the objects (Gonzalez *et al.*, 2008; Umbaugh *et al.*, 2023).

ADDITIVE WHITE GAUSSIAN NOISE

Additive White Gaussian Noise (AWGN) is based on a Gaussian noise model that allows random simulation of pixels with uniformly distributed values in an image. These values can be modeled in a histogram with a Gaussian function. Due to its mathematical low complexity and tractability, the Gaussian noise model is frequently used in various practical applications (Gonzalez *et al.*, 2008; Bovik, 2009). This distribution can be described through its probability density function,

$$f(x) = (2\pi\sigma^2)^{-1/2} e^{-(x-\mu)^2/2\sigma^2}. \quad (2)$$

This equation is determined by the mean μ and the variance σ^2 of a random variable x. This type of noise is difficult to eliminate completely in digital images given its characteristics.

IMAGE CONTRAST

Grayscale image contrast can be defined as the intensity difference between the highest and lowest gray levels in the image. In other words, if the image has an appreciable number of pixels with a high dynamic range, it is likely to display high contrast in its content. On the contrary, an image with a low dynamic range tends to have a washed-out or dull gray appearance (Gonzalez *et al.*, 2008; Russ, 2016). In Fig. 1, two

examples of grayscale synthetic images are illustrated with an AWGN noise vector ($\mu = 0$, $\sigma^2 = 0.001$). Here, it is showed different contrast between background (outer square) and object (inner square). The image in Fig. 1a shows an image with a better grayscale contrast difference between the background and the object. Otherwise, the image in Fig. 1b shows a low contrast image between the background and the object.
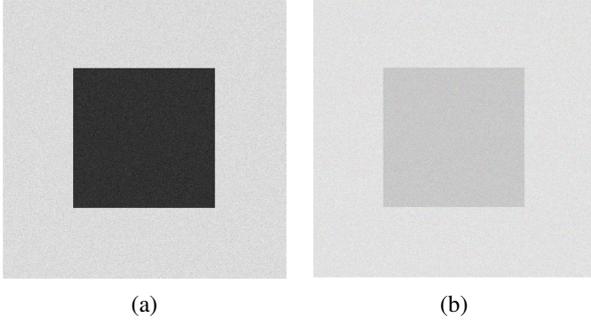

(a)   (b)

Fig. 1: *Synthetic grayscale images with AWGN. (a) High-contrast image. (b) Low-contrast image.*

OTSU'S METHOD

Otsu's method is one of the best threshold selection methods for real images in terms of uniformity and shape measures (Resma *et al.*, 2018; Lu *et al.*, 2017; Lei *et al.*, 2019; Liu *et al.*, 2015). However, Otsu's method employs a deep search to evaluate the criteria and to expand the within-class variation. As the number of classes in an image increases, so does the time constraint for multithreshold selection. The probabilities $p(i)$ of two classes separated by a threshold t are represented by ω_0 and ω_1. The means for each class can be defined as

$$\mu_0(t) = \sum_{i=0}^{t-1} \frac{ip(i)}{\omega_0(t)}, \qquad (3)$$

$$\mu_1(t) = \sum_{i=t}^{L-1} \frac{ip(i)}{\omega_1(t)}, \qquad (4)$$

where the finite range $[0, 256]$ of i ranges from 0 to $t - 1$ and from t to $L - 1$ for μ_0 and μ_1 respectively. This is evaluated for each $p(i)$, which represents the probability of occurrence of the range values of i in the image histogram. The total average intensity of the image is defined as

$$\mu_T = \sum_{i=0}^{L-1} ip(i). \qquad (5)$$

$$\omega_0.\mu_0 + \omega_1.\mu_1 = \mu_T. \qquad (6)$$

$$\omega_0 + \omega_1 = 1. \qquad (7)$$

By using discriminant analysis, the between-class variance of a thresholded image can be calculated as

$$\sigma_B^2 = \omega_0.(\mu_0 - \mu_T)^2 + \omega_1.(\mu_1 - \mu_T)^2. \qquad (8)$$

So, to calculate the optimal threshold value that maximizes the variance between classes for image segmentation, it would be as follows

$$t^* = \arg\max_t \{\sigma_B^2(t)\} \qquad 0 \leq t < L - 1. \quad (9)$$

SIMILARITY METRICS

Since a visual or graphic comparison is insufficient to evaluate the differences between images, two statistical metrics, Dice and Jaccard, are proposed to measure the results. These similarity metrics are commonly used in different works to evaluate segmentation performance (Zou *et al.*, 2004; Taha *et al.*, 2015; Chung *et al.*, 2019).

Dice

The Dice Similarity Coefficient (DSC) is one of the most widely used metrics in the validation of binary segmentation (Taha *et al.*, 2015; Zou *et al.*, 2004). It allows to compare the similarity of resulting binary images A with a true reference image B. The DSC is calculated by the following equation

$$DSC(A, B) = \frac{2|A \cap B|}{|A| + |B|}, \qquad (10)$$

where the symbol $\cap$ denotes the conjunction of two sets and $|\cdot|$ represents the cardinality of a set.

Jaccard

The Jaccard Similarity Index (JSI), in digital images, compares the pixels of two sets $[A, B]$ to see which subset of pixels $[\varphi, \phi]$ are shared and which are different (Kosub, 2019). Fundamentally, it is a measure of similarity for the two data sets, with a percentage range from 0 to 100 and normalized from 0 to 1. The more similar are two populations of pixels corresponding to A and B, the Jaccard index is closer to 1. The equation (11) is used to calculate the Jaccard index (Taha *et al.*, 2015; Kosub, 2019).

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}. \qquad (11)$$

Here the symbols $\cap$ and $\cup$ represent the conjunction and the union of two sets respectively. Whereas the symbol $|\cdot|$ denotes the cardinality of each set.

METHODOLOGY

In this section a detailed description about the employed methodology is presented based on two principal stages Image Acquisition and Experimental Procedure. The database in Image Acquisition was obtained from fruits and vegetables easily accessible in our environment. While, in Experimental Procedure was proposed two processing tests by employing the raw images and adding AWGN noise.

IMAGE ACQUISITION

The images to carry out the experimental tests were selected from our database consisting of 320 RGB images of size 1536×2048 pixels. The selected images were those with the greatest lighting problems. The selected products can be appreciated in Fig. 2 and their commercial names are white onion, caribbean chili, yellow apple, potato, yellow bell pepper, xcatik chili, and pear in that order.
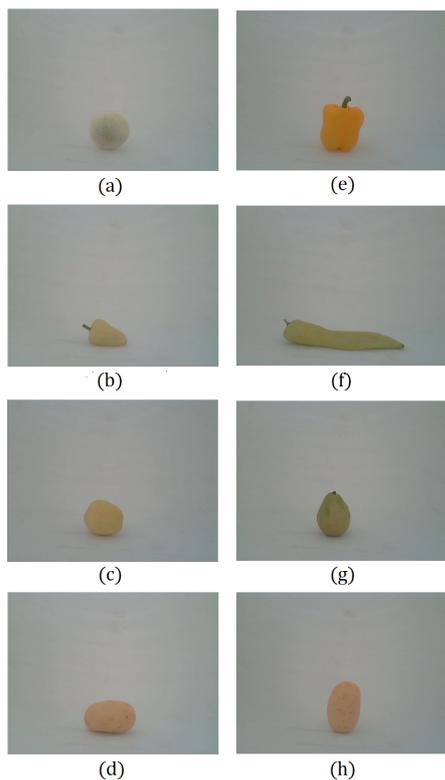

(a)   (e)

(b)   (f)

(c)   (g)

(d)   (h)

Fig. 2: *RGB images of samples used. (a) white onion, (b) caribbean chili, (b) yellow apple, (d) potato 1, (e) yellow bell pepper, (f) xcatik chili, (g) pear, and (h) potato 2.*

These images were captured in a natural lighting environment using an expanded polystyrene booth of size 50 cm × 80 cm × 50 cm, corresponding to the length, width, and height, respectively. A high-definition Logitech C270 webcam with a USB

connection and a video capture resolution of 720 pixels and 30 frames per second was used. A maximum effective resolution of 3 Mega Pixels. The webcam was placed horizontally 50 cm from the object, as illustrated in Fig. 3a.
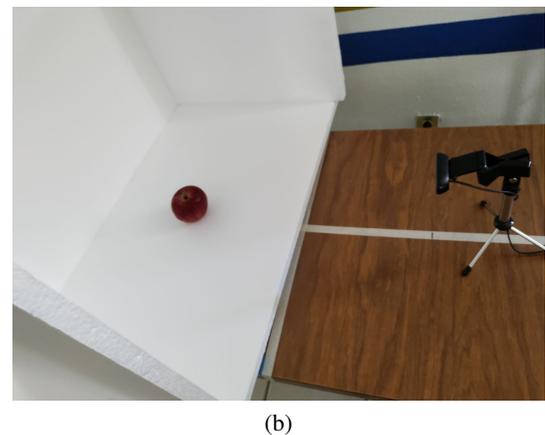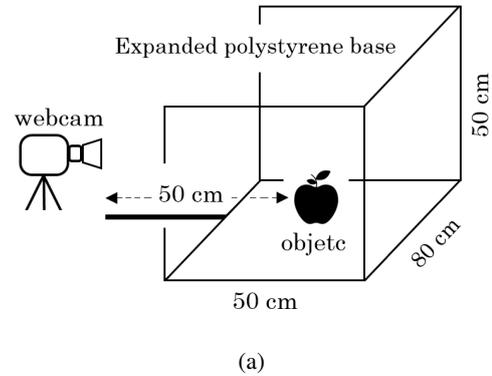

(a)

(b)

Fig. 3: *Schematic diagram of image capture using a webcam, natural lighting, and an expanded polystyrene base. b) An actual picture of the test image capture scenario with the characteristics specified in the previous diagram.*

Images were processed and analyzed using Matlab R2018a software running on Windows 10 Pro 64-bit operating system. An HP workstation with an Intel(R) - Xeon(R) E3-1226 v3 @ 3.30 GHz processor was used. Fig. 3b shows the image of the real scenario in which the test images were obtained. This image was captured with a mobile phone camera. Several samples were obtained with different conditions, however, the images obtained with internal natural lighting and the closed laboratory were used for this experiment. This caused the dividing line at the bottom of the cabin in the images obtained not to be perceived.

EXPERIMENTAL PROCEDURE

The experimental procedure was developed with two tests by processing the original images (I_m) and

noisy images ($I_m + AWGN$), these tests are labeled as e_1 and e_2, respectively. For the case of the e_2 experimental test, it was considered to add AWGN noise to the selected RGB images to test the robustness of the proposed strategy. A general methodology is shown in the diagram of Fig. 4.
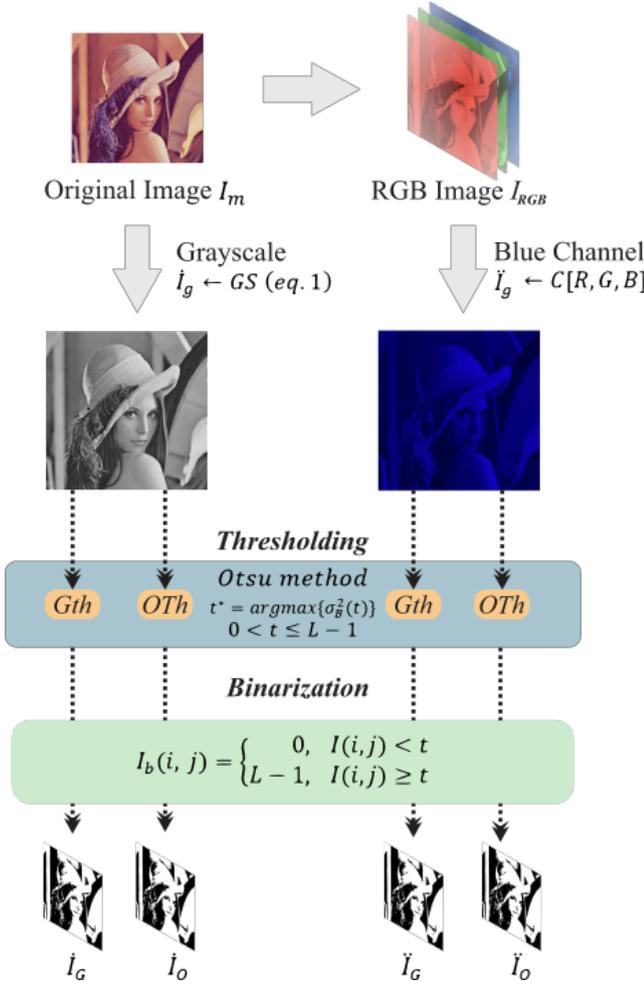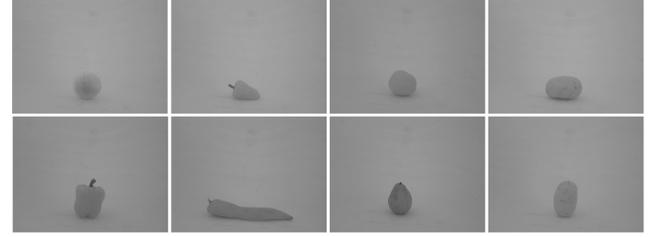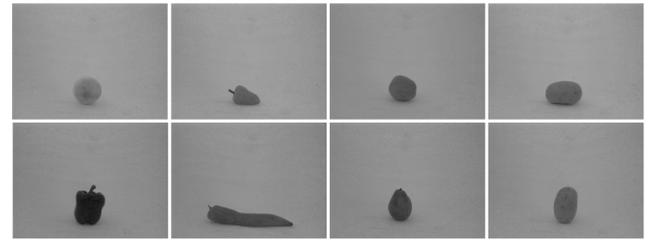

Fig. 4: *Diagrams of the experimental procedure.*

In each experiment, e_1 and e_2, a processing algorithm is applied where four resulting binary images are obtained ($\dot{I}_G$, $\dot{I}_O$, $\ddot{I}_G$ and $\ddot{I}_O$). In the processing algorithm, sketched in Fig. 4, two intensity images are obtained from the selected RGB images. The first image $\dot{I}_g$ was obtained by using the Matlab command `rgb2gray`, this command is based on the equation (1). The second image, expressed as $\ddot{I}_g$, is obtained by selecting the Blue channel from the channel separation of the RGB images. This channel is selected because the Blue component presents a greater contrast in gray intensity between background and object. This fact suggests a greater separation between two pixel subsets φ and ϕ. Therefore, an accurate threshold t can be calculated to obtain a binary image with better

segmentation. The images obtained from $\dot{I}_g$ and $\ddot{I}_g$ can be seen in Fig. 5.

(a) $\dot{I}_g$ image set.

(b) $\ddot{I}_g$ image set.

Fig. 5: $\dot{I}_g$ and $\ddot{I}_g$ images used for experimental tests e_1.

Later, the threshold values of $\dot{I}_g$ and $\ddot{I}_g$ images are calculated by using the commands `graythresh` (*GTh*) and `otsuthresh` (*OTh*). Here, *GTh* calculates a global threshold t from the values in the grayscale of the image. Meanwhile, *OTh* calculates a global threshold t from count calculations in the histogram. Finally, the images are binarized with the threshold t obtained in the previous step.

The processing algorithm is applied in the same way to the images obtained in the e_2 experimental test. However, the variables obtained from this include the subscript $_n$ to note the difference. For example, grayscale images obtained by equation (1) and RGB channel selection can be expressed as $\dot{I}_{gn}$ and $\ddot{I}_{gn}$, respectively. These were not shown in this work due to the similarity with the images in Fig. 5. In a similar way to the experimental test e1, four binary images were obtained following the procedure described in Fig. 4. For the evaluation and comparison of the results obtained, the ground truth images shown in Fig. 6 were used. These samples were manually segmented and approved by the experience of different professionals dedicated to digital image processing, such as Morales-Mendoza *et al.* (2012), Lopez-Ramirez *et al.* (2020), and Gonzalez-Lee *et al.* (2021).
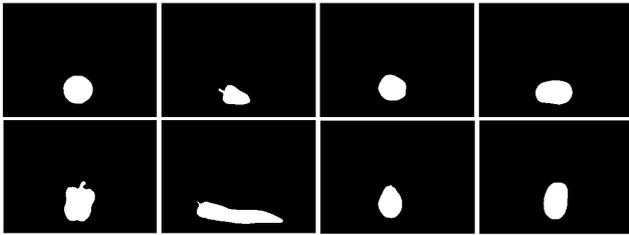
Fig. 6: *Ground truth images.*

RESULTS AND DISCUSSION

FIRST EXPERIMENT

The first section of results to corroborate the improvement segmentation of fruits and vegetables with our proposal is described below. The Fig. 7a shows the images obtained of $\dot{I}_G$ and $\dot{I}_O$ from test e_1. These images were obtained by applying the variants of the Otsu threshold technique GTh and OTh to the images shown in Fig. 5. Some samples of the row of images in Fig. 7a labeled with $\dot{I}_G$ exhibit a poor segmentation. In other words, it is observed that most were severely affected in the background and object by lighting noise. Therefore, the segmentation is visually inaccurate. The least affected images were the bell pepper, xkatik chili and pear. These samples only present small shadows of pixels that are not part of the object. For its part, the row of images of $\dot{I}_O$ were also mostly affected by undesired pixels between the background and object. The images of the onion and Caribbean chili were the least affected. The results of images $\ddot{I}_G$ and $\ddot{I}_O$ are showed in Fig. 7b from the images $\ddot{I}_g$ displayed in Fig. 5. It can be seen that the row of images of $\ddot{I}_G$ and $\ddot{I}_O$ show problems with extra pixels in the background, except for the onion that preserves a minimum proportion of black points in both cases. However, in all images small shadows of pixels can be observed. The results observed in Fig. 7b show a considerable improvement with respect to the results in Fig. 7a, where the GTh and OTh variants were applied. At first sight, there is a better quality in the segmentation of the object with respect to the background.

SECOND EXPERIMENT

Now, a second stage is realized following the same procedure described in Fig. 4 with the RGB images disturbed with AWGN noise with a $\mu = 0$ and variance $\sigma = 0.0002$. The results of the experimental e_2 are shown in Fig. 8. The processed images $\dot{I}_{Gn}$ and $\dot{I}_{On}$ based on the conversion $\dot{I}_{gn}$ are illustrated in Fig. 8a. Here, it can be seen that most of the images were additionally affected by AWGN and illumination noise. In the $\dot{I}_{Gn}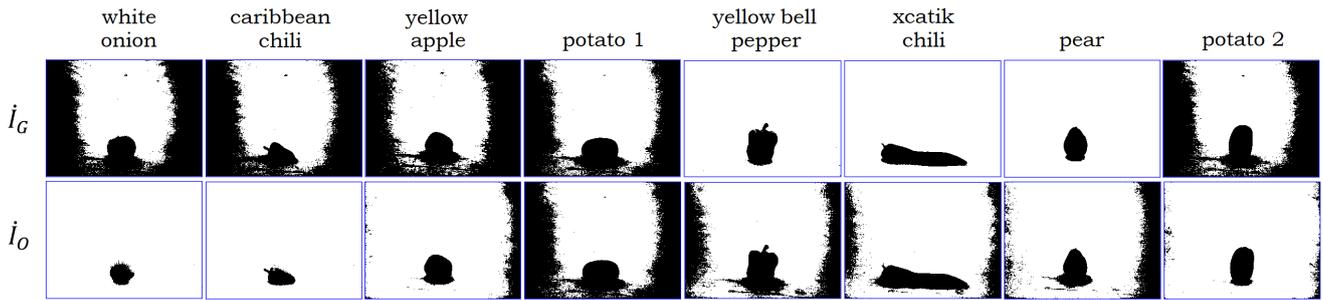$ row of images, bell pepper, xkatik chili, and pear were the least affected in the background. For the row of images in $\dot{I}_{On}$, the onion was less affected in the background, but it was impacted by the loss of pixels that define the object. Fig. 8b shows the results of the $\ddot{I}_{Gn}$ and $\ddot{I}_{On}$ images obtained by applying GTh and OTh to the $\ddot{I}_{gn}$ images. In the row of $\ddot{I}_{Gn}$, it can be seen that most of the images were unperturbed in the segmented background. Only the onion image was highly affected in the background and object due to AWGN and lighting noise. In the case of the row of $\ddot{I}_{On}$ images, the onion, and the caribbean chili were affected with pixels that contaminate the background of the image and some small shadows that are not part of the object. Fig. 8b shows a similar behavior to the previous case, e_1, where these results obtained are better compared to those shown in Fig. 8a.
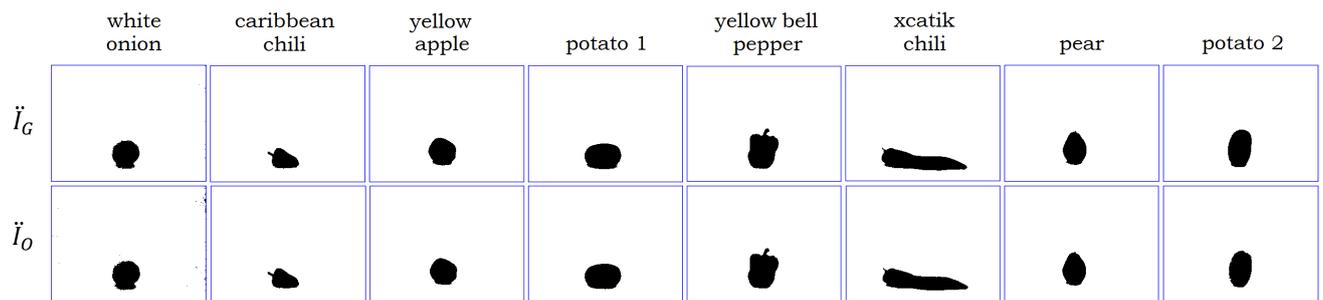
IMPROVEMENT EVALUATION

To measure the results of the experimental tests e_1 and e_2, it is required to use some metrics such as Dice and Jaccard to evaluate them quantitatively. For this purpose, ground truth images shown in Fig. 6 were used. In these measurements of similarity of results using Dice and Jaccard metrics, four comparative tables of experimental test results e_1 and e_2 were obtained.

For this analysis, the symbol ρ represents the value of improvement that exists between the absolute difference of the error percentage values obtained from the JSI and DSC results using the formula $\% \ error = \left| (V_o - V_r)/V_r \right| * 100$. Where V_o is the value obtained from JSI or DSC and V_r is the valid value of 1 as the true reference. The improvement values ρ represent a comparative improvement difference of results of $\ddot{I}_G$ vs. $\dot{I}_G$, $\dot{I}_O$ vs. $\dot{I}_O$, $\ddot{I}_{Gn}$ vs. $\dot{I}_{Gn}$, and $\ddot{I}_{On}$ vs. $\dot{I}_{On}$.

Table 1 exhibits the comparison of results obtained by using Dice and Jaccard metrics based on the images $\dot{I}_G$ and $\ddot{I}_G$, which were obtained from the experimental test e_1. In this table, it is observed that the results obtained from the $\ddot{I}_G$ column are vastly better than those obtained from the $\dot{I}_G$ column in most cases. This is because they are closer to the value of one and are interpreted as good segmentation. In addition, there is a difference in percentage improvement ρ_1 greater than 73% in most of the results, except for bell pepper, chile xkatik, and pear, which present very small differences in percentage improvement ρ_1 that are too small, less than 7%. However, they are good results because their metric values are very close to one.

Fig. 7: *Results obtained from the e_1 experimental test applying GTh and OTh on the $\dot{I}_g$ and $\ddot{I}_g$ images. (a) $\dot{I}_G$ and $\dot{I}_O$ images, (b) $\ddot{I}_G$ and $\ddot{I}_O$ images.*
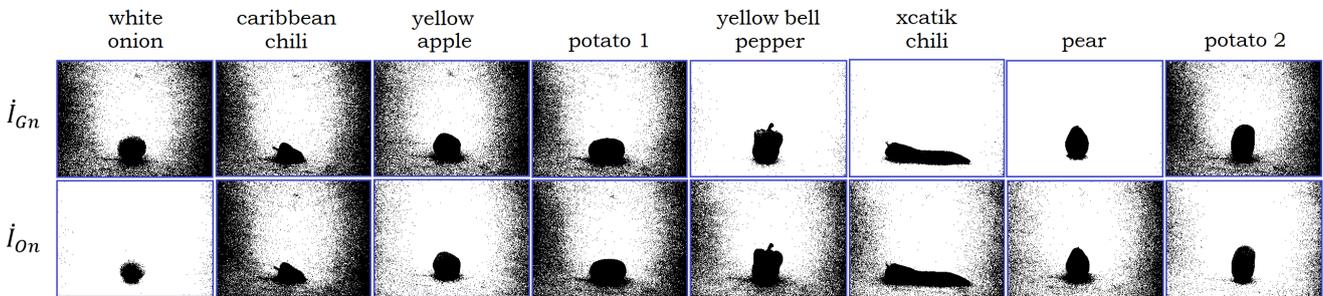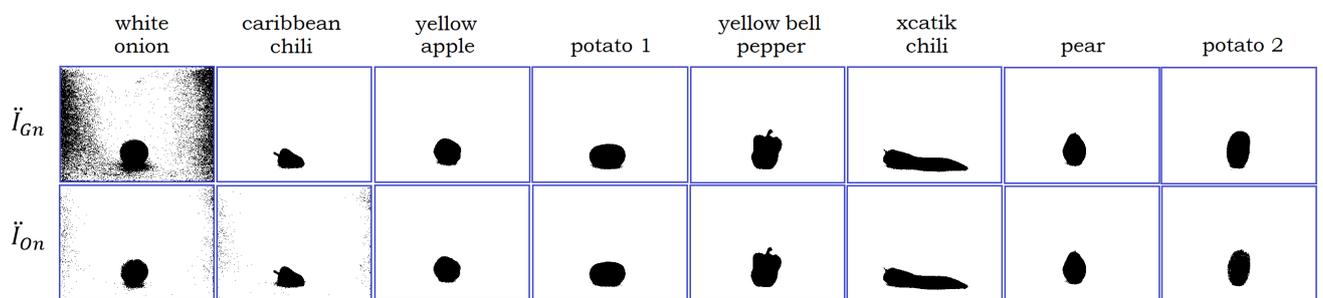

Fig. 8: *Results obtained from the e_2 experimental test applying GTh and OTh on the $\dot{I}_{gn}$ and $\ddot{I}_{gn}$ images. (a) $\dot{I}_{Gn}$ and $\dot{I}_{On}$ images, (b) $\ddot{I}_{Gn}$ and $\ddot{I}_{On}$ images.*

Table 1: *Comparison of results obtained by the JSI and DSC metrics applied to the rows of binary images resulting $\dot{I}_G$ and $\ddot{I}_G$ of Fig. 7 in the experimental test e_1. The ρ_1 results represent the improvement in segmentation obtained between $\ddot{I}_G$ and $\dot{I}_G$ based on the error percentages of the JSI and DSC values.*

Images	Metrics	$\dot{I}_G$	$\ddot{I}_G$	$\rho_1(\%)$
White onion	DSC	0.1135	0.8830	**76.95**
	JSI	0.0602	0.7905	**73.03**
Caribbean chili	DSC	0.0818	0.9713	**88.95**
	JSI	0.0427	0.9442	**90.15**
Apple	DSC	0.1138	0.9665	**85.27**
	JSI	0.0603	0.9352	**87.49**
Potato 1	DSC	0.1588	0.9551	**79.63**
	JSI	0.0863	0.9141	**82.78**
Bell pepper	DSC	0.9565	0.9935	03.70
	JSI	0.9166	0.9871	07.05
Xkatik chilli	DSC	0.9361	0.9668	03.07
	JSI	0.8798	0.9358	05.60
Pear	DSC	0.9497	0.9770	02.73
	JSI	0.9041	0.9551	05.10
Potato 2	DSC	0.1491	0.9877	**83.86**
	JSC	0.0805	0.9758	**89.53**

In Table 2, the $\dot{I}_O$ and $\ddot{I}_O$ results obtained with the Dice and Jaccard metrics are compared.

Table 2: *Comparison of results obtained by the JSI and DSC metrics applied to the rows of binary images resulting $\dot{I}_O$ and $\ddot{I}_O$ of Fig. 7 in the experimental test e_1. The ρ_2 results represent the improvement in segmentation obtained between $\ddot{I}_O$ and $\dot{I}_O$ based on the error percentages of the JSI and DSC values.*

Images	Metrics	$\dot{I}_O$	$\ddot{I}_O$	$\rho_2(\%)$
White onion	DSC	0.6927	0.8684	17.57
	JSI	0.5299	0.7674	23.75
Caribbean chilli	DSC	0.9054	0.9765	07.11
	JSI	0.8271	0.9540	12.69
Apple	DSC	0.3948	0.9712	**57.64**
	JSI	0.2459	0.9440	**69.81**
Potato 1	DSC	0.1588	0.9592	**80.04**
	JSI	0.0863	0.9215	**83.52**
Bell pepper	DSC	0.2716	0.9929	**72.13**
	JSI	0.1571	0.9858	**82.87**
Xkatik chilli	DSC	0.4425	0.9706	**52.81**
	JSI	0.2841	0.9428	**65.87**
Pear	DSC	0.2667	0.9787	**71.20**
	JSI	0.1539	0.9582	**80.43**
Potato 2	DSC	0.7297	0.9656	23.59
	JSI	0.5744	0.9335	35.91

Here, it is observed that the $\ddot{I}_O$ values are better than the $\dot{I}_O$ results because they almost reach the perfect value. The Caribbean pepper values show the smallest improvement difference ρ_2 using OTh. Onion and potato 2 show an acceptable percentage improvement difference ρ_2, with a mean of 20.58% for Dice and

29.83% for Jaccard. The results for apple and chile xkatik in this table show a ρ_2 improvement difference of more than 50%. The best results are obtained with potato 1, bell pepper and pear, with a difference in improvement ρ_2 greater than 70%. This is reflected in the second row of the images in Figs. 7a and 7b.

For the second experiment e_2, the Table 3 collected the results obtained with the Dice and Jaccard metrics.

Table 3: *Comparison of results obtained by the JSI and DSC metrics applied to the rows of binary images resulting $\dot{I}_{Gn}$ and $\ddot{I}_{Gn}$ of Fig. 8 in the experimental test e_2. The ρ_3 results represent the improvement in segmentation obtained between $\ddot{I}_{Gn}$ and $\dot{I}_{Gn}$ based on the error percentages of the JSI and DSC values.*

Images	Metrics	$\dot{I}_{Gn}$	$\ddot{I}_{Gn}$	$\rho_3(\%)$
White onion	DSC	0.1314	0.2244	09.30
	JSI	0.0703	0.1264	05.61
Caribbean chili	DSC	0.0961	0.9679	**87.18**
	JSI	0.0505	0.9377	**88.72**
Apple	DSC	0.1384	0.9623	**82.39**
	JSI	0.0743	0.9274	**85.31**
Potato 1	DSC	0.1882	0.9521	**76.39**
	JSI	0.1039	0.9086	**80.47**
Bell pepper	DSC	0.8852	0.9915	10.63
	JSI	0.7941	0.9831	18.90
Xkatik chilli	DSC	0.9134	0.9649	05.15
	JSI	0.8406	0.9322	09.16
Pear	DSC	0.9492	0.9802	03.10
	JSI	0.9032	0.9611	05.79
Potato 2	DSC	0.1135	0.9698	**85.63**
	JSI	0.0951	0.9413	**84.62**

As in experimental test e_1 and the results of the previous tables, it is observed that the results obtained using DSC and JSI to the $\ddot{I}_{Gn}$ images are significantly better than the results of the $\dot{I}_{Gn}$ images. Now, there is a percentage improvement difference ρ_3 of more than 75% in the Caribbean chili, apple, potato1 and potato 2. For xkatik chili and pear, the differences in improvement ρ_3 are small, but better in segmentation quality because they present values close to one in the DSC and JSI results in $\ddot{I}_{Gn}$ and $\dot{I}_{Gn}$. The onion case presents a very small ρ_3 improvement difference, but of low quality in the segmentation due to its values close to zero. This example corroborated that most of the AWGN noise and bad lighting noise in the onion image could not be eliminated. This can be seen in the row $\ddot{I}_{Gn}$ and $\dot{I}_{Gn}$ images in Fig. 8.

For its part, Table 4 shows that the results obtained from $\ddot{I}_{On}$ are better than those processed with $\dot{I}_{On}$. Most of the results of the Dice and Jaccard metrics for $\ddot{I}_{On}$ are near to 0.9, which is considered a good segmentation. Opposite of this, the results of $\dot{I}_{On}$ obtained an average of 0.6, which is considered a poor segmentation. About the improvement difference ρ_4, it was achieved a 73%

in the case of potato 1, bell pepper, and pear respect to Dice. In the case of the apple and xkatik chili, a quite acceptable ρ_4 improvement difference of more than 58% was obtained. In the case of potato 2, a ρ_4 computes a 44% enhancement. This result is good because the values in $\ddot{I}_{On}$ present high values, i.e. a good segmentation. In the case of onion, although it presents a small ρ_4 improvement difference, their results are not suitable. The caribbean chilli presents a ρ_4 improvement difference of more than 61%. Despite this, the $\ddot{I}_{On}$ values are regular. This causes that an amount of pixels remain contaminating the background and the object, as shown in the row of the respective images $\dot{I}_{On}$ and $\ddot{I}_{On}$ of Fig. 8.

Table 4: *Comparison of results obtained by the JSI and DSC metrics applied to the rows of binary images resulting $\dot{I}_{On}$ and $\ddot{I}_{On}$ of Fig. 8 in the experimental test e_2. The ρ_4 results represent the improvement in segmentation obtained between $\ddot{I}_{On}$ and $\dot{I}_{On}$ based on the error percentages of the JSI and DSC values.*

Images	Metrics	$\dot{I}_{On}$	$\ddot{I}_{On}$	$\rho_4(\%)$
White onion	DSC	0.6522	0.7616	10.94
	JSI	0.4839	0.6150	13.11
Caribbean chili	DSC	0.0859	0.7942	**70.83**
	JSI	0.0449	0.6586	**61.37**
Apple	DSC	0.3014	0.9698	**66.84**
	JSI	0.1774	0.9414	**76.40**
Potato 1	DSC	0.1680	0.9577	**78.97**
	JSI	0.0917	0.9188	**82.71**
Bell pepper	DSC	0.2594	0.9917	**73.23**
	JSI	0.1490	0.9836	**83.46**
Xkatik chilli	DSC	0.3822	0.9696	**58.74**
	JSI	0.2362	0.9410	**70.48**
Pear	DSC	0.2196	0.9814	**74.18**
	JSI	0.1233	0.9635	**84.02**
Potato 2	DSC	0.4970	0.9435	**44.65**
	JSI	0.3307	0.8931	**56.24**

Table 5 shows the best results obtained from experiments e_1 and e_2. In the comparison of the results of $\ddot{I}_G$ and $\ddot{I}_O$ of the experimental test e_1, small differences can be observed between the best results equally distributed. Although most Dice and Jaccard results achieved an average value greater than 0.9, in the particular case of the onion, it presents the lowest values in the comparison of the experimental test e_1. In the case of the experimental test e_2, it can be observed that most of the results are good for Dice and Jaccard showing a small advantage for the $\ddot{I}_{On}$ results. As in the test e_1, the worst results were obtained by processing the onion image. Here, the metrics of $\ddot{I}_{On}$ were regular and the images obtained for $\ddot{I}_{Gn}$ were lousy.

Table 5: *Comparison between the best results of the two experimental tests, e_1 and e_2. First, the comparison of the $\ddot{I}_G$ vs. $\ddot{I}_O$ results using the DSC and JSI metrics. Then, the second comparison, $\ddot{I}_{Gn}$ vs. $\ddot{I}_{On}$ with the same metrics.*

Images	Metrics	e_1		e_2	
		$\ddot{I}_G$	$\ddot{I}_O$	$\ddot{I}_{Gn}$	$\ddot{I}_{On}$
White onion	DSC	**0.8830**	0.8684	0.2244	**0.7616**
	JSI	**0.7905**	0.7674	0.1264	**0.6150**
Caribbean chili	DSC	**0.9713**	0.9765	**0.9679**	0.7942
	JSI	**0.9442**	0.9540	**0.9377**	0.6586
Apple	DSC	0.9665	**0.9712**	0.9623	**0.9698**
	JSI	0.9352	**0.9440**	0.9274	**0.9414**
Potato 1	DSC	0.9551	**0.9592**	0.9521	**0.9577**
	JSI	0.9141	**0.9215**	0.9086	**0.9188**
Bell pepper	DSC	**0.9935**	0.9929	0.9915	**0.9917**
	JSI	**0.9871**	0.9858	0.9831	**0.9836**
Xcatik chilli	DSC	0.9668	**0.9706**	0.9649	**0.9696**
	JSI	0.9358	**0.9428**	0.9322	**0.9410**
Pear	DSC	0.9770	**0.9787**	0.9802	**0.9814**
	JSI	0.9551	**0.9582**	0.9611	**0.9635**
Potato 2	DSC	**0.9877**	0.9656	**0.9698**	0.9435
	JSI	**0.9758**	0.9335	**0.9413**	0.8931

CONCLUSIONS

The results of the experimental tests proposed show that the images with higher contrast can obtain a better segmentation despite lighting problems. This argument is supported with graphic and numerical evidence. Most of the results show an average percentage improvement difference greater than 45.5% in the two experimental tests. A small percentage improvement difference in the two experimental tests does not exhibit if the result is good or bad. This rating depends on the range in which the values obtained by the Dice and Jaccard metrics are located.

Based on the metrics of Dice and Jaccard in the test e_1, both modalities of Otsu's method obtained the same number of best results. That is to say, the best result of segmentation varies according to the modality of Otsu's method and the characteristics of the image in the test e_1. In the $\ddot{I}_{Gn}$ and $\ddot{I}_{On}$ results of the e_2 test, clearly the Dice and Jaccard metrics show that the otsuthresh modality obtained the most improvement results.

The results in general were satisfactory using this experimental procedure, but it would be required to work with different database images and with AWGN variations. Furthermore, for future work, it is intended to compare the proposed strategy with modern segmentation algorithms such as those based on artificial intelligence.

ACKNOWLEDGMENTS

REFERENCES

Alegre E, Pajares G, De la Escalera A (2016). Conceptos y métodos en visión por computador. España: Grupo de Visión del CEA, 11–28.

Bataineh B, Abdullah SNHS, Omar K (2017). Adaptive binarization method for degraded document images based on surface contrast variation. Pattern Anal Applic 20, 639–52.

Belan PA, de Macedo RAG, Alves WAL, Santana JCC, Araújo SA (2020). Machine vision system for quality inspection of beans. Int J Adv Manuf Technol 111, 3421–35.

Bovik AC (2009). The essential guide to image processing. Academic Press, 2009.

Bradley D, Roth G (2007). Adaptive thresholding using the integral image. Jou Graph Tools 12(2), 13–21.

Chávez OV, Flores–Troncoso J, Minjares JUM, Reyna RO, Sánchez EG, Reyna RO (2022). Image segmentation of Capsicum annuum chili with lighting problems using the Otsu method. Stud in Engin and Exact Scienc 3(4), 560–73.

Cheremkhin PA, Kurbatova EA (2019). Comparative appraisal of global and local thresholding methods for binarisation of off-axis digital holograms. Opt Lasr Enging 115, 119–30.

Chung NC, Miasojedow BZ, Startek M, Gambin A. (2019). Jaccard/Tanimoto similarity test and estimation methods for biological presence–absence data. BMC Bioinfo 20(15), 1–11.

Fan P, Lang G, Guo P, Liu Z, Yang F, Yan B, Lei X (2021). Multi-feature patch-based segmentation technique in the gray-centered RGB color space for improved apple target recognition. Agriculture 11(3), 273.

Gatos B, Pratikakis I, Perantonis SJ (2006). Adaptive degraded document image binarization. Pattern Recogn 39(3), 317–27.

Goh TY, Basah SN, Yazid H, Safar MJA, Saad FSA (2018). Performance analysis of image thresholding: Otsu technique. Measurem, 114, 298–307.

Gonzalez RC, Woods RE (2008). Digital image processing, N.J.: Prentice Hall, 91–320.

Gonzalez-Lee M, Vazquez-Leal H, Morales-Mendoza LJ, Nakano-Miyatake M, Perez-Meana H, Laguna-Camacho JR (2021). Statistical assessment of discrimination capabilities of a fractional calculus based image watermarking system for Gaussian watermarks. Entropy 23(2), 255.

Kosub S (2019). A note on the triangle inequality for the Jaccard distance. Pattern Recogn Lett 120, 36-38.

Lopez-Ramirez M, Ledesma-Carrillo LM, Garcia-Guevara FM, Munoz-Minjares JU, Cabal-Yepez E, Villalobos-Pina FJ (2020). Automatic early broken-rotor-bar detection and classification using Otsu segmentation. IEEE Access 8, 112624–32.

Lei B, Fan J (2019). Image thresholding segmentation method based on minimum square rough entropy. App Soft Comp 84, 105687.

Liu L, Yang N, Lan J, Li J (2015). Image segmentation based on gray stretch and threshold algorithm. Optik 126(6), 626-29.

Lu Y, Lu R (2017). Histogram-based automatic thresholding for bruise detection of apples by structured-illumination reflectance imaging. Biosys Enginng 160, 30–41.

Morales-Mendoza LJ, Vázquez-Bautista RF, Morales-Mendoza E, Semenovich Shmaliy Y (2012). Speckle noise reduction in ultrasound imaging using the key points in low degree unbiased fir filters. Comp. y Sists. 16(3), 287-95.

Mukhiddinov M, Muminov A, Cho J (2022). Improved Classification Approach for Fruits and Vegetables Freshness Based on Deep Learning. Sensors 22(21), 8192.

Ng HF (2006). Automatic thresholding for defect detection. Pattern Recogn Lett 27(14), 1644–49.

Niblack W (1986). An introduction to digital image processing. Cliffs, N.J.: Prentice Hall, 115-16.

Otsu N (1979). A threshold selection method from gray-level histograms. IEEE Trans SMC 9(1), 62–6.

Resma KPB, Nair MS (2021). Multilevel thresholding for image segmentation using Krill Herd Optimization algorithm. J kSU-Comp and Infor Sciens 33(5), 528–41.

Russ JC (2006). The Image Processing Handbook, 5th edition. USA, CRC, 31-149.

Sahoo PK, Soltani S, Wong AKC (1988). A survey of thresholding techniques. Comput Vis, Graph, and Image Process 41(2), 233–60.

Sauvola J, Pietikäinen M (2000). Adaptive document image binarization. Patt recogn 33(2), 225–36.

Sezgin M, Sankur BL (2004). Survey over image thresholding techniques and quantitative performance evaluation. JEM, 13(1), 146–68.

Sha C, Hou J, Cui H(2016). A robust 2D Otsu's thresholding method in image segmentation. Jou of VC and IR 41, 339–51.

Singh, TR, Roy S, Singh OI, Sinam T, Singh KM (2011). A new local adaptive thresholding technique in binarization. Int Jou of Comput Scienc Iss 8, 6(2), 271–77.

Song SB, Liu JF, Ni HY, Cao XL, Pu H, Huang BX (2020). A new automatic thresholding algorithm for unimodal gray-level distribution images by using the gray gradient information. Jo of Petrol Scien and Enging 190, 107074.

Sundararajan D (2017). Digital image processing: a signal processing and algorithmic approach. Canada: Springer (1), 1-58.

Taha AA, Hanbury A (2015). Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. BMC Medic Imagn 15(1), 1–28.

Umbaugh SE (2023). Digital image processing and analysis: computer vision and image analysis. 4th ed., CRC Press, 1-35.

Yang W, Cai L, Wu F (2020). Image segmentation based on gray level and local relative entropy two dimensional histogram. PLoS ONE 15(3), e0229651.

Yuan XC, Wu LS, Peng Q (2015). An improved Otsu method using the weighted object variance for defect detection. Appld Surfac Sciec 349, 472–84.

Zheng X, Ye H, Tang Y (2017). Image bi-level thresholding based on gray level-local variance histogram. Entropy 19(5), 191.

Zou KH, Warfield SK, Bharatha A, Tempany CM, Kaus MR, Haker SJ, Wells WM, Jolesz FA, Kikinis R (2004). Statistical validation of image segmentation quality based on a spatial overlap index[1]: scientific reports. Academs Radiol 11(2), 178−89.