ISSN 2007-9737

Comparative Evaluation of Segmentation Techniques on Maize Seeds

Juan Augusto Campos-Leal[1,*], Eduardo Diaz-Gaxiola[1], Alejandro Lizarraga-Sarabia[1], Jair Cervantes-Canales[2], Farid García-Lamont[2], Ines Fernando Vega-Lopez[1], Arturo Yee-Rendon[1]

[1] Universidad Autónoma de Sinaloa, México

[2] Universidad Autónoma del Estado de México, Mexico

{juan.campos, eduardogaxiola, ifvega, arturo.yee}@uas.edu.mx, a.lizarraga19@info.uas.edu.mx, {jcervantesc, fgarcial}@uaemex.mx

Abstract. In this study, we performed a comparative analysis of various segmentation techniques on a dataset of digital images depicting maize seeds. The dataset includes categories representing different quality levels of seed batches, labeled as worst, bad, average, good, and excellent. The objective of this study was to identify the most accurate combination of segmentation techniques and color spaces for maize seed segmentation. To achieve this, we systematically compared various segmentation techniques, including Otsu, Watershed, and *K*-means clustering, across different color spaces such as RGB, HSV, and grayscale. We evaluated the performance of these combinations using metrics like Intersection over Union (IoU), Mean Square Error (MSE), Pixel Accuracy (PA), Kappa (κ) coefficient, and F-1 Score in images from the Maize Seed dataset. The results showed that the *K*-means segmentation technique in the HSV color space yielded the best results across all evaluated metrics. In particular, using the IoU metric in the HSV color space, *K*-means achieved mean values of 0.98, 0.97, 0.95, 0.91, and 0.91 for the categories excellent, good, average, bad, and worst, respectively.

Keywords. Segmentation techniques, maize seed, machine learning.

1 Introduction

Maize is one of the most vital crops in Mexico, contributing significantly to the country's economy and holding substantial social and cultural importance. In 2021, Mexico was the seventh largest producer of maize globally, with a production of 27 million tons [15]. The state of Sinaloa plays a key role in this output, contributing 22% of the national production [2].

Despite its importance, maize cultivation faces challenges such as pests and unfavorable climate conditions, which can lead to production difficulties and economic losses. These factors underscore the importance of assessing maize quality rigorously for local, national, and international trade purposes.

One effective approach to evaluate maize quality is through seed inspection, which typically involves visually examining seeds sampled from a batch to assess their overall condition and suitability for planting or consumption. However, this visual assessment can be subjective and prone to human error, especially when distinguishing subtle differences in seed quality. Factors such as seed size, shape, color, and surface texture play crucial roles in determining their viability and nutritional value.

The process of seed inspection can significantly benefit from leveraging computer vision as a powerful tool. Computer vision addresses the challenge of designing artificial visual systems capable of interpreting our three-dimensional real world. A fundamental step in this process is

identifying the object of interest and separating it from the surrounding background in an image. This process, known as image segmentation [6], involves grouping pixels in a digital image to form regions.

Selecting a proper segmentation technique is critical because of the wide variety of techniques available for segmenting objects in digital images. These techniques range from simple thresholding and region-based approaches to advanced techniques such as clustering algorithms. It is essential to select a technique that matches the unique characteristics of the objects being segmented and fulfills the specific requirements of the application.

In this work, we propose a comparative study of computer vision techniques for segmenting maize seeds. We used seed batch images ranging from excellent to poor quality. We evaluated three segmentation techniques Otsu, Watershed, and *K*-means across three color spaces: RGB, HSV, and grayscale. To assess the techniques used in this study, we employed five evaluation metrics: Intersection over Union (IoU), Mean Square Error (MSE), and Pixel Accuracy (PA), Kappa (κ) coefficient, and F-1 Score. Our findings showed that the *K*-means algorithm performed best in the HSV color space. Overall, the results obtained by the segmentation techniques in the HSV color space surpassed those in the other color spaces.

Additionally, we considered the computational efficiency and robustness of each segmentation technique, ensuring that our findings apply to real-world scenarios. This comprehensive evaluation provides insights into the most effective approaches for accurate and reliable maize seed segmentation, which can significantly enhance agricultural practices, such as seed sorting and quality control.

The remainder of this paper is organized as follows: Section 2 provides a summary of relevant work. Section 3 describes the materials and methods used in this study, including the color spaces, segmentation techniques, and dataset. Section 4 details the computing environment, evaluation metrics used and presents the results of our experimental evaluation. Finally, Section 5 presents our conclusions and recommendations for further research.

2 Related Work

In this section, we present the works related to our comparative evaluation and the applied segmentation on different seeds. We review prior research addressing similar challenges in agricultural imaging, focusing on diverse segmentation techniques applied to different types of seeds. Our analysis of these studies aims to clarify our methodology and contribute to the discussion on improving seed analysis through computer vision techniques.

In 2019, Tan et al. [18] introduced an algorithm to separate and count touching rice grains. Their approach integrates the watershed algorithm, an enhanced corner point detection method, and a neural network classifier. To improve the contrast intensity of grayscale images and reduce noise, the authors applied a wavelet transform and a Gaussian filter. This preprocessing step mitigates over-segmentation typically caused by the watershed algorithm. Given the variation in appearance and corner point characteristics across regions with different grain quantities, an artificial neural network classifier was used to categorize under-segmented regions into five classes: single grain, two grains, three grains, four grains, and more than four grains. The method was tested on samples consisting of 300 rice grains per variety, achieving an average accuracy of 94.63% compared to manual counting.

In 2020, Abbas et al. [1] proposed a methodology for detecting and calculating the ratio of the affected area on maize leaves. This methodology comprises three main stages: (a) classifying the image using *K*-means clustering, (b) segmenting it through color thresholding, and (c) estimating the affected area. The approach was evaluated across different color spaces: RGB, HSV, YCbCr, and L*a*b*. The authors built a dataset with images from the International Maize and Wheat Improvement Center, featuring symptoms such as Eyespot sores, Macrospora lesions, and Hyalothyridium lesions. The results indicated that color thresholding was more efficient

than *K*-means clustering in determining the injury ratio, particularly in the HSV color space. In contrast, the RGB color model produced the least accurate results.

In 2021, Bai et al. [3] developed an automatic supplemental seeding device using computer vision techniques to correct seeding errors in plug seeders, replacing the need for manual reseeding. They proposed a voting mechanism-based method for detecting missing sweet corn seeds in a moving plug tray. The methodology included median filtering, morphological processing, contour extraction, and seed candidate selection, applied to both RGB and HSV color spaces. To enhance the accuracy of seed recognition, a voting mechanism combined the detection results from both color spaces, considering contour areas and the locations of seed candidates. The device achieved approximately 99% accuracy in detecting sweet corn seeds.

In 2022, Liang et al. [13] introduced a segmentation method for handling complex-touching grains of various species, shapes, and sizes. The approach involved three specialized algorithms to segment non-touching, simple-touching, and complex-touching grains in captured images, enhancing segmentation accuracy. The *K*-means algorithm was used to extract non-touching grains, while a layered watershed algorithm was developed for segmenting simple-touching grains. For complex-touching grains, an innovative splitting line detection algorithm was devised, incorporating the determination of concave corner point sets, line detection, and line extension for precise segmentation. The method was evaluated on 100 top-view images for each grain type, including peanuts, corn, soybeans, red beans, mung beans, unhusked rice, wheat, black rice, rice, and rapeseed, all captured with a top camera and white backlight. Testing on 1,000 images comprising 640,000 grains demonstrated an average measurement accuracy of 99.65%.

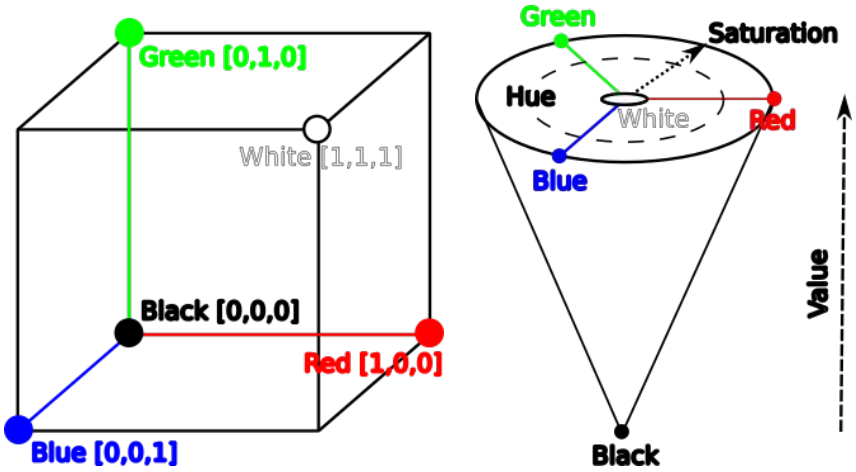
In 2023, Song et al. [16] introduced a seed-quality evaluation method using an enhanced Inception-ResNet network for assessing corn seeds of varying qualities. Initially, images of multiple corn seeds were captured using a multispectral camera to create a dataset of individual seed images. The dataset comprised 50 groups, including 982 corn seeds in total—458 in good condition and 524 damaged. These groups represented seeds with good appearance, defective grains, and a mix of both conditions. The adaptive thresholding, morphological open operation and, the watershed algorithm were used during image segmentation. An attention mechanism was also incorporated to enhance the network's feature learning and optimize the extraction of image information representing seed appearance quality. The method demonstrated robust overall performance in detecting corn seed appearance quality, achieving an average detection accuracy of 96.03%, precision of 96.27%, recall of 96.03%, and an F-1 Score of 96.15%.

In 2024, Rai et al. [14] developed a method for segmenting and detecting Northern Leaf Blight (NLB) in maize crops using advanced deep learning techniques. The authors evaluated three models: Attention U-Net, Res U-Net, and Plain U-Net. They used the Maize Disease dataset [19], which includes over 18,222 RGB images of maize leaves captured by various devices—a handheld camera, a boom-mounted camera, and a small UAV at a 6-meter altitude. These images were taken from field trials of maize inoculated with the pathogen Setosphaeria turcica, responsible for NLB. Expert annotations identified more than 105,735 NLB lesions within the dataset. The models achieved Intersection over Union (IoU) scores of 72.41%, 70.91%, and 51.95% for Attention U-Net, Res U-Net, and Plain U-Net, respectively. The proposed method attained an average pixel-wise F-1 Score of 85.23%, a diseased segmentation accuracy of 98.97%, and a Dice coefficient (DC) of 81.39% for disease spot segmentation.

3 Materials and Methods

In this section, we present a brief definition of color spaces used in this comparative evaluation. We mention the segmentation techniques evaluated. Furthermore, we present the dataset used for this study.

ISSN 2007-9737



(a) RGB color space. (b) HSV color space.

Fig. 1. Geometrical representations of RGB space and HSV space in color. Images adapted from Garcia-Lamont et al. [10].

3.1 Color Spaces

A digital image, informally defined as a sequence of pixels forming patterns that determine a representation, is typically encoded in a computer using a specific color space.

One widely used example is the RGB color space, which is used for transmitting, representing, and storing color images in both analog and digital devices [8].

A color space is designed to facilitate the representation of colors in a three-dimensional coordinate system. It can also be used to define a subset of this system, where each color corresponds to a unique point. Many of the color spaces that are used are specific to hardware, like printers and monitors, or to color-manipulating applications, like animation graphic design.

Furthermore, the representation and processing of colors in images depend on the choice of color space.

The color spaces used in this analysis are RGB (Red, Green, Blue) and HSV (Hue, Saturation, Value), the most commonly used in the literature according to Garcia-Lamont et al. [10].

3.1.1 RBG Space

With its origins in television technology, the RGB color space is regarded as the basic color representation for digital devices. In this space, spectral components of red (R), green (G), and blue (B) collectively represent every color. The intensity of these components determines both the hue and brightness of resulting colors. The inclusion of each component while keeping the color black as the foundation determines the color combination.

The RGB space is shaped like a cube geometrically, as can be seen in Figure 1 (a) with the three color coordinates representing the red, green, and blue components. The values of each component span from 0 to 255, where each point on the cube represents a unique color.

3.1.2 HSV Space

The components hue (h), saturation (s), and value (v) are used to express color in the HSV space. The chromatic characteristic of a pure color is its hue. Value refers to the color's brightness or intensity, whereas saturation measures how the hue is subdued in white light. There are two key components to the HSV space. Value, the intensity component, is separated from the hue

information. The hue and saturation levels simulate how color is perceived by the human eye. Because of these characteristics, the HSV space can be used to create image processing algorithms that take advantage of certain aspects of how people perceive color. Saturation ranges from 0 to 1 in the real domain, hue component from 0 to $2\pi \in \Re$, and value $[0, 255] \in \Re$.

The HSV space is shaped geometrically like a cone, as can be seen in Figure 1 (b) with the saturation and value components represented by the cone's height and radius, respectively. It should be noted that the hue component is undefined for colors like black, white, and gray. Since these colors lack specified Chromaticity, they are regarded as singularities within the HSV color space. The Chromaticity describes the perceived color of an object or light source without considering its brightness level.

3.2 Segmentation Techniques

The process of dividing an image into several segments or regions in order to simplify its representation and improve its analytical value is known as image segmentation. Splitting an image into sections with comparable visual elements, such as color, intensity, texture, or motion, is the aim of segmentation [17].

3.2.1 Otsu's Algorithm

The Otsu method, also known as Otsu's thresholding or Otsu's algorithm, is a popular technique used for automatic image thresholding. Using this technique, the ideal threshold value to divide an image's pixels into foreground and background classes is automatically determined. By maximizing the difference in intensity values between the two classes, the approach seeks to identify a threshold that maximizes the degree of differentiation between the foreground and background [6]. The Otsu method is described in Algorithm 1.

Algorithm 1 Otsu's Thresholding Algorithm

Require: Grayscale image I
Ensure: Optimal threshold T
1: Compute the histogram H of image I
2: Compute the total number of pixels N and total mean m
3: Initialize $w_B \leftarrow 0$, $w_F \leftarrow 0$, $sum_B \leftarrow 0$, $varMax \leftarrow 0$, $T \leftarrow 0$
4: **for** each intensity level t in H **do**
5: Update background weight $w_B \leftarrow w_B + H[t]$
6: **if** $w_B = 0$ **then**
7: **continue**
8: **end if**
9: Update foreground weight $w_F \leftarrow N - w_B$
10: **if** $w_F = 0$ **then**
11: **break**
12: **end if**
13: Update background mean $m_B \leftarrow sum_B / w_B$
14: Update foreground mean $m_F \leftarrow (m - sum_B)/w_F$
15: Compute between-class variance $varBetween \leftarrow w_B \times w_F \times (m_B - m_F)^2$
16: **if** $varBetween > varMax$ **then**
17: Update maximum variance $varMax \leftarrow varBetween$
18: Update optimal threshold $T \leftarrow t$
19: **end if**
20: Update cumulative sum $sum_B \leftarrow sum_B + t \times H[t]$
21: **end for**
22: **return** T

3.2.2 Watershed Algorithm

Its foundation is the idea of simulating the image as a topographic surface, with elevations represented by pixel intensities. The fundamental principle of watershed segmentation is to treat the image intensity gradient as a topographic surface, with low intensity gradients corresponding to flat areas and high intensity gradients corresponding to areas of steep climb. The markers that the computer initially inserts in the image serve as the segmentation's beginning points. These marks can be created automatically or manually in accordance with predetermined standards. Subsequently, the algorithm models a flooding

process in which each marker point releases water onto the topographic surface. The peaks and valleys in the intensity gradient create basins that are flooded when the water level rises. Eventually, when the water from different basins meets at the watershed lines, these lines become the boundaries between different segmented regions [11]. The watershed segmentation algorithm is presented in 2.

Algorithm 2 Watershed Segmentation Algorithm

Require: Image I, Marker image M, Connectivity structure C
Ensure: Segmented image S
Convert I to grayscale and compute the gradient magnitude image G
Apply markers to I to get the marker image M
Initialize the priority queue Q with boundary pixels of markers in M
while priority queue Q is not empty **do**
    Pop pixel p with lowest gradient value from Q
    for each neighboring pixel n of p according to connectivity C **do**
        if n is unmarked in M **then**
            Assign n to the same region as p in M
            Add n to the priority queue Q
        end if
    end for
end while
Use marker image M to segment the original image I
return Segmented image S

3.2.3 *K*-means for Color Quantization

It is a technique used in image processing to diminish the number of colors while upholding visual fidelity. The objective is to represent the image with fewer colors, termed "clusters", thus reducing memory usage for storage or transmission without compromising image quality [7]. The algorithm begins by randomly selecting *K* initial cluster centers within the color space, where *K* denotes the desired number of colors in the output image. Subsequently, each pixel in the input image is assigned to the nearest cluster center based on color similarity, effectively grouping the pixels into *K* clusters. Following pixel assignment, the algorithm computes the mean color (centroid) for each cluster, representing the average color of its assigned pixels. The cluster assignments and centroid calculation steps iterate until they stabilize. Upon convergence, each pixel in the input image is substituted with the color of the nearest centroid, resulting in the final quantified image with *K* colors [11]. The *K*-means for color quantization is described in Algorithm 3.

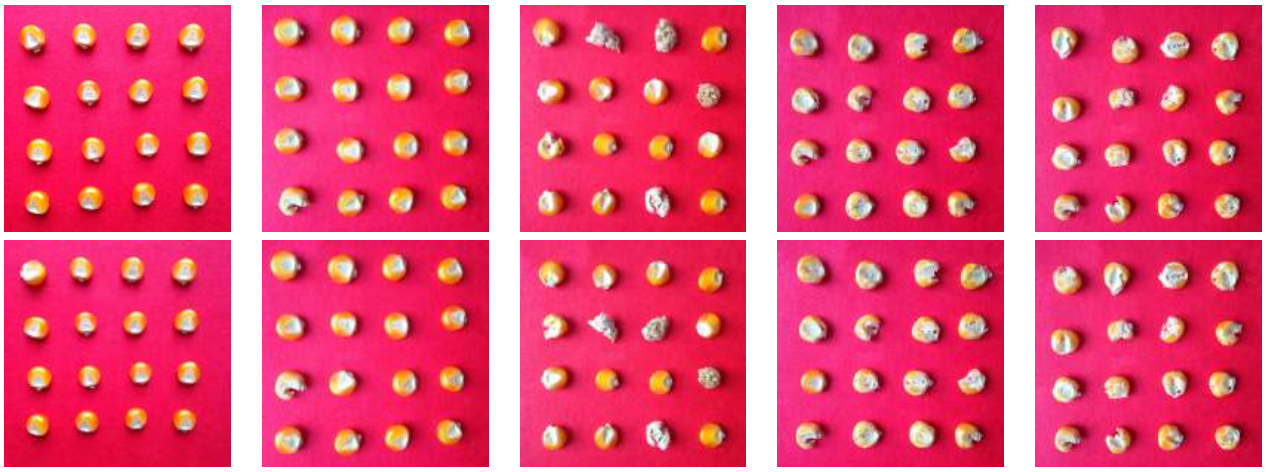
Algorithm 3 *K*-means clustering for color quantization

Require: Image I, Number of clusters k, Maximum iterations $maxIter$
Ensure: Segmented image S
Convert I to a 2D array of pixels P
Initialize k centroids C randomly
for $i = 1$ to $maxIter$ **do**
    Assign each pixel in P to the nearest centroid in C
    Update each centroid in C to the mean of assigned pixels
end for
Assign each pixel in P the color of its centroid in C
Reshape P back to the original image dimensions to obtain S
return S

3.3 Maize Seed Dataset

The Maize Seed dataset [5] can be downloaded from Kaggle [4]. This dataset consists of five categories of seeds, determined by the state of a seed batch. The five categories are Excellent, Good, Average, Bad, and Worst. Figure 2 shows sample images for each category. Each category in the dataset is determined by the percentage of damaged seeds in the seed batch.

To determine a seed batch as an Excellent category, the percentage of damaged seeds is 0 percent. For the Good category, the percentage ranges from 1 to 39 percent. The Average category includes percentages from 40 to 60 percent. Percentages between 61 and 99 percent



(a) (b) (c) (d) (e)

Fig. 2. Categories in the Maize Seed dataset. (a) Excellent (b) Good (c) Average (d) Bad (e) Worst

fall into the Bad category. Lastly, the Worst category is characterized by 100 percent damaged seeds. The dataset comprises 2,500 training images and 500 testing images distributed among these five categories.

3.3.1 Preprocessing

We manually segmented the images to generate ground truth masks for the dataset. Each resulting mask is a binary image, with white denoting the seed's region of interest and black indicating the background. Figure 3 shows some examples of ground truth masks for the images in the Maize Seed dataset. Those examples masks corresponding to the images in Figure 2.

4 Results and Discussion

This section describes the results of our comparative evaluation. Firstly, we describe the computing environment and the used evaluation metrics. Secondly, we report the results obtained by the segmentation techniques used on RGB and HSV color space. We present the outcomes applied to a maize seed segmentation task. Our analysis was conducted on the Maize Seed dataset: five distinct categories was evaluated on the segmentation task.

4.1 Computational Environment and Metrics Descriptions

The experiments were conducted on a computer equipped with an Intel(R) Xeon(R) Silver 4210R processor, 32 GB of RAM, Ubuntu 22.04 served as the operating system, accompanied with Python version 3.8.0. The OpenCV-Python 4.6.0 library was used. To evaluate the segmentation techniques, we used Intersection over Union (IoU), Mean Square Error (MSE), Pixel Accuracy (PA), Kappa (κ) coefficient, and F-1 Score metrics. These indices are defined as follows.

Intersection over Union (IoU) measures the overlap between the predicted segmentation mask and the ground truth segmentation mask of an object in an image. It is calculated by dividing the area of intersection (the overlapping region) between the predicted and ground truth segmentation masks by the area of union (the combined region) of the two segmentation masks. The IoU is defined by Equation 1:

$$\text{IoU} = \frac{\text{Area of Intersection}}{\text{Area of Union}}. \tag{1}$$

The IoU value ranges from 0 to 1, where IoU = 0 indicates no overlap between the predicted and ground truth segmentation masks, and IoU = 1 indicates a perfect overlap, meaning the predicted

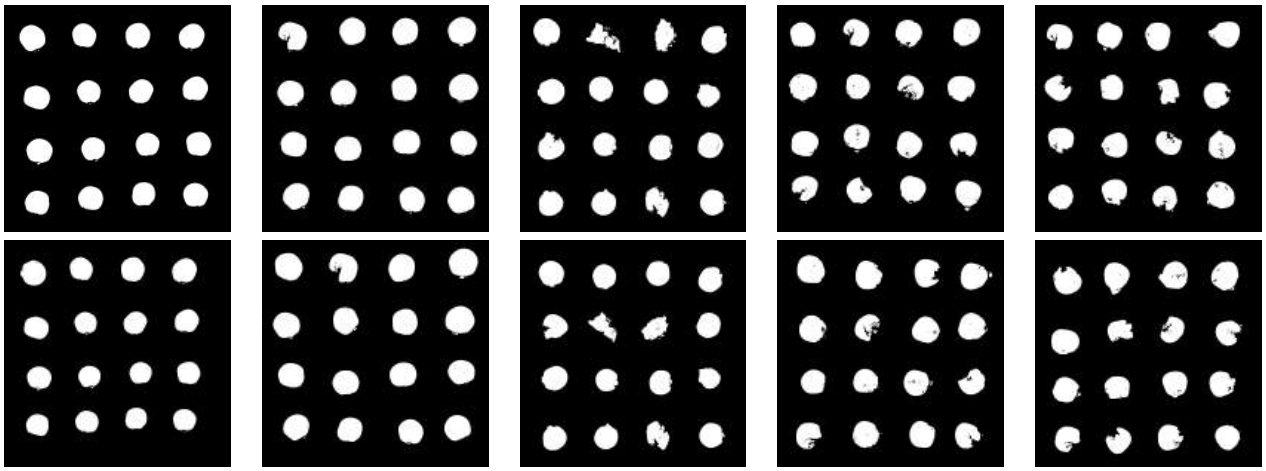


Fig. 3. Examples of ground truth masks in the Maize Seed dataset. (a) Excellent (b) Good (c) Average (d) Bad (e) Worst

segmentation mask exactly matches the ground truth segmentation mask.

In the context of image segmentation, Mean Square Error (MSE) measures the average squared difference between the pixel values of the predicted segmentation mask and those of the ground truth segmentation mask across all pixels in the image. The MSE is defined by Equation 2:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(p_i - g_i)^2, \tag{2}$$

N is the total number of pixels in the image. p_i is the predicted pixel value at position i. g_i is the ground truth pixel value at position i. A lower MSE value indicates closer agreement between the predicted segmentation and the ground truth, suggesting higher accuracy of the segmentation algorithm.

Pixel accuracy (PA), also referred to as pixel-level accuracy, is a metric used to assess the performance of image segmentation tasks. It quantifies the percentage of accurately classified pixels in an image relative to the total number of pixels. Pixel accuracy is computed by dividing the number of correctly classified pixels by the total number of pixels in the image. The PA is defined by Equation 3:

$$\text{PA} = \frac{\text{Number of Correctly Classified Pixels}}{\text{Total Number of Pixels}}. \tag{3}$$

A high value in pixel accuracy indicates that a large percentage of pixels in the segmented image are classified correctly compared to the ground truth.

The Kappa index (or Cohen's Kappa coefficient) [12] [9] is a statistical measure of inter-rater agreement for qualitative (categorical) items. It is used to evaluate the agreement between two raters or segmentation methods beyond chance:

$$\kappa = \frac{Ac - \rho}{1 - \rho}, \tag{4}$$

where:

Ac is Accuracy, it given by:

$$\text{Ac} = \frac{TP + TN}{TP + TN + FP + FN}. \tag{5}$$

The ρ is given by:

$$\rho = \frac{(TP + FN) \times (TP + FP) + (TN + FN) \times (TN + FP)}{(TP + TN + FP + FN)^2}, \tag{6}$$

where:

Table 1. Results using the Intersection over Union (IoU) metric

	RGB		HSV		Gray
	Watershed	***K*-means**	**Watershed**	***K*-means**	**Otsu**
Avg.	0.9000	0.7265	0.9204	**0.9504**	0.7434
Med.	0.9101	0.7357	0.9280	**0.9663**	0.7437
Max.	0.9774	0.9668	0.9905	**0.9910**	0.9085
Min.	0.6628	0.3295	0.6857	**0.7237**	0.5446

Table 2. Results using the Mean Square Error (MSE) metric

	RGB		HSV		Gray
	Watershed	***K*-means**	**Watershed**	***K*-means**	**Otsu**
Avg.	0.0157	0.0417	0.0125	**0.0076**	0.0388
Med.	0.0139	0.0409	0.0112	**0.0052**	0.0388
Max.	0.0573	0.1720	0.0537	**0.0395**	0.0714
Min.	0.0030	0.0052	0.0012	**0.0012**	0.0137

TP is the True Positives, FP is the False Positives, TN is the True Negatives, and FN is the False Negatives.

Based on Equation 4, the accuracy level (AL) of the segmented images is is categorized according to the κ value as follows:

$$\text{AL} = \begin{cases} P, & \kappa \leq 0.2 \\ R, & 0.2 < \kappa \leq 0.4 \\ G, & 0.4 < \kappa \leq 0.6 \\ VG, & 0.6 < \kappa \leq 0.8 \\ E, & \kappa > 0.8 \end{cases}, \tag{7}$$

where P is Poor, R is Reasonable, G is Good, VG is Very Good, and E correspond to Excellent.

The Precision, Recall, F-1 Score metrics are also used to evaluate the performance.

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{8}$$

$$\text{Recall} = \frac{TP}{TP + FN}, \tag{9}$$

$$\text{F-1 Score} = 2 \times \frac{Precision \cdot Recall}{Precision + Recall}. \tag{10}$$

Precision measures the proportion of correctly predicted positive observations to the total predicted positives. Recall measures the proportion of correctly predicted positive observations to all observations in the actual class. The F-1 Score is the harmonic mean of Precision and Recall. The F-1 Score provides a single value that balances both Precision and Recall.

4.2 Experimental Results

In this section, we present the results of segmentation techniques using the previously mentioned performance metrics. Each table details the average, median, maximum, and minimum values derived from analyzing images in RGB, HSV, and grayscale color spaces. These metrics collectively assess the segmentation methods across diverse color representations, emphasizing their efficacy and resilience in varied image contexts.

In Table 1, the Intersection over Union (IoU) metric is reported for both RGB and HSV color spaces, while the Otsu method required assessment specifically in the grayscale color space. A value closer to 1 indicates better performance according to IoU definition. Notably, the *K*-means algorithm in the HSV color space achieved the best average value. Additionally, *K*-means in the HSV color space attained the highest maximum value compared to other segmentation techniques, including RGB and grayscale color spaces. Table 2 shows the results for the Mean Squared Error (MSE) metric, where a value closer to 0 indicates better performance.

Table 3. Results using the Pixel Accuracy (PA) metric

	RGB		HSV		Gray
	Watershed	***K*-means**	**Watershed**	***K*-means**	**Otsu**
Avg.	0.9842	0.9582	0.9874	**0.9923**	0.9611
Med.	0.9860	0.9590	0.9887	**0.9947**	0.9611
Max.	0.9969	0.9947	0.9987	**0.9987**	0.9862
Min.	0.9426	0.8279	0.9462	**0.9604**	0.9285

Table 4. Results using the Kappa (ρ) coefficient

	RGB				HSV				Gray	
	Watershed		***K*-means**		**Watershed**		***K*-means**		**Otsu**	
Avg.	0.9374	E	0.8159	E	0.9506	E	**0.9696**	**E**	0.8300	E
Med.	0.9446	E	0.8243	E	0.9561	E	**0.9796**	**E**	0.8310	E
Max.	0.9867	E	0.9799	E	0.9945	E	**0.9946**	**E**	0.9435	E
Min.	0.7655	VG	0.4131	G	0.7871	VG	**0.8173**	**E**	0.6708	VG

It is evident that the *K*-means algorithm in the HSV color space yielded the best results, closely followed by the watershed algorithm in the same color space.

Table 3 corresponds to the results for the Pixel Accuracy (PA) metric. Although the majority of values are close to 1, indicating high accuracy, the *K*-means algorithm in the HSV color space achieved the best performance.

Table 4 reports the results using the Kappa (ρ) coefficient. According to this metric, we noted that the results are established in general as Excellent, and just noted some Very Good, and Good for the minimum values. Similar to other metrics used, the *K*-means algorithm in HSV reached the best results as well.

In Table 5, the F-1 Score metric is used to report the results of the segmentation. The results showed that *K*-means algorithm in the HSV color space obtained the best values in F-1 Score.

Table 6 presents the results by categories in the Maize Seed dataset. Each row corresponds to the specific category, and the values represent the average of Intersection over Union (IoU) metric for each category. It is evident from the results that all five categories achieved an IoU higher than 0.91, indicating strong segmentation performance. These high IoU values were primarily attained using the *K*-means technique in the HSV color space, followed by results obtained by the watershed technique in both HSV and RGB color spaces. The *K*-means technique in the RGB color space and results obtained by the Otsu method exhibited slightly lower performance.

4.3 Discussion

Overall, the results reported in Table 1, Table 2, Table 3, Table 4, and Table 5 demonstrate that the superior performance across the used metrics was achieved in the HSV color space. These results consistently surpassed those obtained in the RGB space, indicating that the HSV color space using the *K*-means algorithm is more effective for maize seed segmentation.

Figure 4 shows an example image from the excellent category of the dataset. Figure 4 (a) corresponds to the original image from the dataset, Figure 4 (b) represents the ground truth mask. Figure 4 (c) the prediction of this image using *K*-means in the HSV color space, and Figure 4 (d) depicts the prediction using *K*-means in the RGB color space. Visually, we observe some differences in seed segmentation, which correspond to the variations reported by the metrics used.

We observed consistent performance improvements with the *K*-means algorithm, which consistently outperformed other segmentation techniques such as watershed. In the HSV color space,

Table 5. Results using the $F-1$ Score metric

	RGB		HSV		Gray
	Watershed	***K*-means**	**Watershed**	***K*-means**	**Otsu**
Avg.	0.9466	0.8393	0.9580	**0.9741**	0.8520
Med.	0.9529	0.8477	0.9627	**0.9828**	0.8530
Max.	0.9885	0.9831	0.9952	**0.9954**	0.9520
Min.	0.7972	0.4956	0.8136	**0.8397**	0.7052

Table 6. Results per category using the Intersection over Union (IoU) metric

Categories	RGB		HSV		Gray
	Watershed	***K*-means**	**Watershed**	***K*-means**	**Otsu**
Excellent	0.9358	0.7131	0.9508	**0.9850**	0.7342
Good	0.9255	0.7319	0.9409	**0.9775**	0.7581
Average	0.9038	0.6860	0.9238	**0.9598**	0.7385
Bad	0.8657	0.7392	0.8898	**0.9107**	0.7435
Worst	0.8692	0.7624	0.8967	**0.9190**	0.7427

the watershed technique also showed competitive results but fell slightly behind *K*-means.

Figure 5 shows examples of atypical results during the segmentation process. These examples correspond to the segmentation of two images with *IDs* 359 and 396 from the "Excellent" category in the dataset.

In Figure 5 (a) and Figure 5 (b) examples, segmentation was performed using the *K*-means algorithm in the RGB color space. We observed that the brightness in the original images adversely affected the segmentation process.

Additionally, the image with *ID* 359 also exhibited atypical results when segmented using the watershed algorithm in both color spaces. These examples can be seen in Figure 5 (c) and Figure 5 (d).

Finally, we noted that the Otsu method performed poorly in grayscale, and both *K*-means and watershed algorithms did not yield satisfactory results in the RGB color space.

This suggests that the choice of segmentation technique and color space significantly impacts segmentation accuracy, with the HSV color space and *K*-means proving to be the most effective combination for this study.

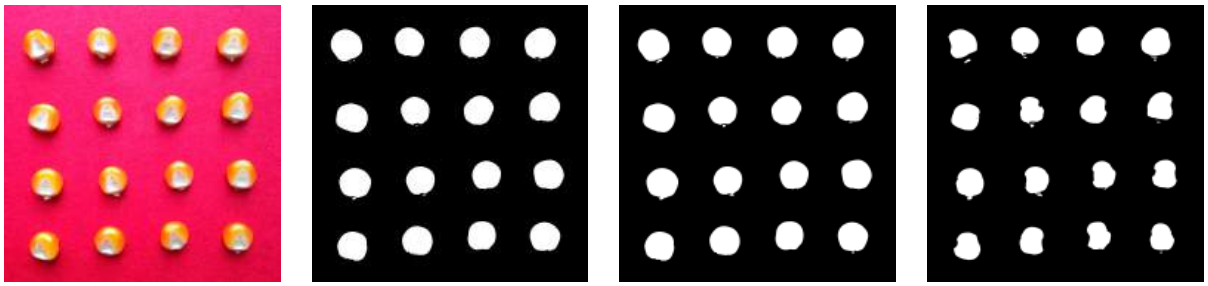
5 Conclusions

In this study, we systematically compared different segmentation techniques across various color spaces to find the most effective combination for maize seed segmentation. Using the *K*-means algorithm in the HSV color space consistently produced superior results across all evaluation metrics (IoU, MSE, PA, κ, and F-1). This approach outperformed both RGB and grayscale color spaces, particularly when compared to the Otsu method. Furthermore, we noted that the HSV color space significantly enhanced the differentiation between parts of maize seeds, thereby improving segmentation accuracy.

As future work, we aim to extend our evaluation to include other types of seeds commonly found in Mexico. This expansion will allow us to further our findings across different agricultural contexts and enhance the applicability of our segmentation approach.

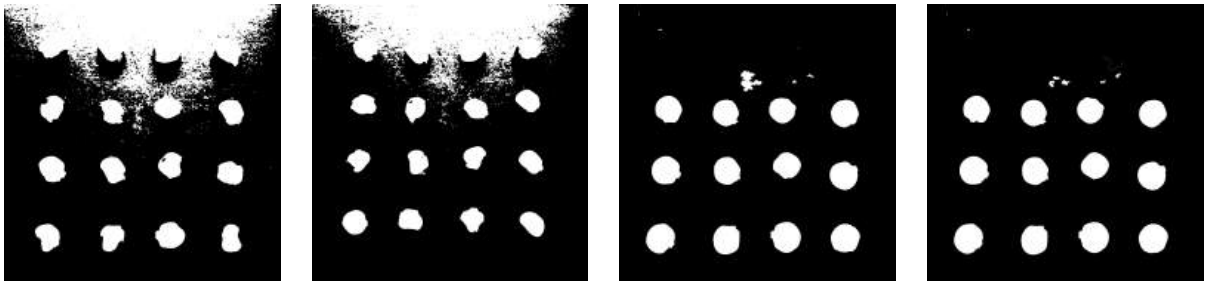
Acknowledgments

The authors express their sincere gratitude to the Mexican Council of Science, Humanities, and Technology (CONAHCYT) for awarding a scholarship to the first author. Finally, they extend



(a) (b) (c) (d)

Fig. 4. Original image, ground truth mask, and segmented images using *K*-means in HSV and RGB color spaces. Example image from the excellent category. (a) Original (b) Ground truth (c) *K*-means in HSV (d) *K*-means in RGB



(a) ID: 359 (b) ID: 396 (c) ID: 359 (d) ID: 359

Fig. 5. Four atypical results of segmentation. Example images from the excellent category. (a) and (b) Examples using *K*-means algorithm in RGB. (c) Example using watershed algorithm in RGB. (d) Example using watershed algorithm in HSV

their appreciation to the Universidad Autónoma de Sinaloa (UAS) for its continued support.

References

1. **Abbas, A. H., Mirza, N. M., Qassir, S. A., Abbas, L. H. (2020).** Maize Leaf Images Segmentation Using Color Threshold and K-means Clustering Methods to Identify the Percentage of the Affected Areas. Proceedings of the IOP Conference Series: Materials Science and Engineering, Baghdad, Iraq, December 16-17, IOP Publishing, Vol. 745, No. 1, pp. 012048.
2. **Agencia de Servicios a la Comercialización y Desarrollo de Mercados Agropecuarios (2024).** Maíz Grano Cultivo Representativo de México. https://www.gob.mx/aserca/articulos/maiz-grano-cultivo-representativo-de-mexico. Accessed: 2024-03-23.
3. **Bai, J., Hao, F., Cheng, G., Li, C. (2021).** Machine Vision-based supplemental Seeding Device for Plug Seedling of Sweet Corn. Comput. Electron. Agric., Vol. 188, pp. 106345. DOI: 10.1016/J.COMPAG.2021.106345.
4. **Bhurtel, M. (2018).** Maize Seed. https://www.kaggle.com/dsv/173263. DOI: 10.34740/KAGGLE/DSV/173263.
5. **Bhurtel, M., Shrestha, J., Lama, N., Bhattarai, S., Uprety, A., Guragain, M. K. (2019).** Deep Learning Based Seed Quality Tester. Science, Engineering and Technology (SET) Conference.

6. **Bhuyan, M. K. (2019).** Computer Vision and Image Processing: Fundamentals and Applications. CRC Press.
7. **Bishop, C. M. (2007).** Pattern Recognition and Machine Learning, 5th Edition. Information science and statistics. Springer.
8. **Burger, W., Burge, M. J. (2016).** Digital Image Processing - An Algorithmic Introduction Using Java, Second Edition. Texts in Computer Science. Springer. DOI: 10.1007/978-1-4471-6684-9.
9. **García-Lamont, F., Alvarado, M., López-Chau, A., Cervantes, J. (2022).** Efficient Nucleus Segmentation of White Blood Cells Mimicking the Human Perception of Color. Color Research & Application, Vol. 47, No. 3, pp. 657–675. DOI: 10.1002/col.22752.
10. **García-Lamont, F., Cervantes, J., López Chau, A., Rodríguez-Mazahua, L. (2018).** Segmentation of Images by Color Features: A Survey. Neurocomputing, Vol. 292, pp. 1–27. DOI: 10.1016/J.NEUCOM.2018.01.091.
11. **Howse, J., Minichino, J. (2020).** Learning OpenCV 4 Computer Vision with Python 3: Get to Grips with Tools, Techniques, and Algorithms for Computer Vision and Machine Learning. Packt Publishing Ltd.
12. **Landis, J. R., Koch, G. G. (1977).** The Measurement of Observer Agreement for Categorical Data. Biometrics, pp. 159–174.
13. **Liang, N., Sun, S., Yu, J., Taha, M. F., He, Y., Qiu, Z. (2022).** Novel Segmentation Method and Measurement System for various Grains with complex touching. Comput. Electron. Agric., Vol. 202, pp. 107351. DOI: 10.1016/J.COMPAG.2022.107351.
14. **Rai, C. K., Pahuja, R. (2024).** Northern Maize Leaf Blight Disease Detection and Segmentation using deep Convolution Neural Networks. Multim. Tools Appl., Vol. 83, No. 7, pp. 19415–19432. DOI: 10.1007/S11042-023-16398-3.
15. **Secretaría de Agricultura y Desarrollo Rural (2024).** Maíz Grano Cultivo Representativo de México. https://www.gob.mx/agricultura/articulos/maiz-cultivo-de-mexico. Accessed: 2024-03-23.
16. **Song, C., Peng, B., Wang, H., Zhou, Y., Sun, L., Suo, X., Fan, X. (2023).** Maize Seed Appearance Quality Assessment based on Improved Inception-ResNet. Frontiers in Plant Science, Vol. 14, pp. 1249989.
17. **Szeliski, R. (2022).** Computer Vision: Algorithms and Applications. Springer Nature.
18. **Tan, S., Ma, X., Mai, Z., Qi, L., Wang, Y. (2019).** Segmentation and Counting Algorithm for Touching Hybrid Rice Grains. Comput. Electron. Agric., Vol. 162, pp. 493–504. DOI: 10.1016/J.COMPAG.2019.04.030.
19. **Wiesner-Hanks, T., Stewart, E. L., Kaczmar, N., DeChant, C., Wu, H., Nelson, R. J., Lipson, H., Gore, M. A. (2018).** Image Set for Deep Learning: Field Images of Maize Annotated with Disease Symptoms. BMC research notes, Vol. 11, pp. 1–3.

Article received on 17/07/2024; accepted on 15/02/2025.
**Corresponding author is Juan Augusto Campos-Leal.*