

Clasificación automática de tipos de semilla de quinua a través de descriptores de color

John C. Montalvo-Samame*, Victor A. Tuesta-Monteza, Juan C. Arcila-Diaz, Heber I. Mejia-Cabrera

Escuela de Ingeniería de Sistemas, Facultad Ingeniería Arquitectura Urbanismo, Universidad Señor de Sipán
Pimentel, Chiclayo – Perú

{*msamamej, vtuesta, diarcilaju, hmejia} @crece.uss.edu.pe

Keywords: Clasificación de semillas, *Chenopodium quinoa*, descriptores de color.

Abstract

Los agricultores de quinua para obtener cosechas optimas deben seleccionar de manera adecuada sus semillas a cultivar sin que se mezclen con otras variedades. La investigación actual se centra en la clasificación automática de tres tipos de semillas de quinua (Sacaca, Pasankalla y Salcedo) utilizando descriptores de color. Después de la adquisición de imágenes de semilla de quinua, se les asigna el filtro Gaussianblur para corregir y cuantificar el color en las imágenes permitiendo resaltar las diferencias entre las características de cada tipo de semilla de quinua. Las imágenes suavizadas se asignan al proceso de segmentación utilizando el método de Otsu para extraer después las características y realizar el entrenamiento de los clasificadores. Para realizar la clasificación de las semillas de quinua se utilizó SVM mediante el análisis lineal pixel a pixel. Los resultados de ensayo demuestran que el procedimiento de desarrollo tiene una alta precisión.

1 Introduction

La quinua (*Chenopodium quinoa*) es una planta herbácea originaria de América andina, específicamente de la hoya del Titicaca, entre Perú y Bolivia, lugar donde se encuentra la mayor cantidad de variedades y se cultiva desde hace 7000 años [1].

Las semillas de quinua son pequeños gránulos con diámetros de entre 1.8 y 2.2mm que constituyen un gran valor alimenticio, se considera una importante cosecha de alimentos en los Andes y ha suscitado recientemente interés en los Estados Unidos, Europa y Asia, donde la demanda de quinua se estima en varias miles de toneladas por año por encima del nivel actual de producción [2].

En el Perú existen aproximadamente 18 variedades de quinua, diferenciándose por su color (blanco, amarillo, marrón, negro, plomo, púrpura, cristalino) y por su forma (cónica, esférica, redonda aplanada, lenticular, esférica) [3].

En los andes peruanos los resultados de cosecha según las variedades de quinua dependen mucho del clima donde se cultive, algunas semillas de quinua tienen mejores resultados de cosecha en zonas más altas, es por esto que los agricultores para obtener cosechas optimas deben seleccionar de manera

adecuada sus semillas a cultivar sin que se mezclen con otras variedades, esta tarea de clasificación la pueden realizar los mismos agricultores o las pueden adquirir de empresas agrícolas que venden semillas certificadas según la variedad de quinua, estas empresas ocupan bastante trabajo porque clasifican las semillas de quinua de manera manual pudiendo además tener un margen de error.

Específicamente, en los últimos años se han desarrollado varios trabajos de investigación enfocados al procesamiento de imágenes digitales para la clasificación de semillas [4]. En la investigación [5] se emplearon técnicas de análisis de las características de las texturas de imágenes de trigo para lograr la identificación de nueve variedades de semillas de trigo iraní. En el trabajo desarrollado en [6] se utilizaron imágenes hiperespectrales para desarrollar métodos de clasificación e identificación de semillas de maíz.

Si bien la mayor parte de las investigaciones se centran en cultivos populares como el trigo y el maíz hay una falta de obras centradas en un cultivo actualmente demandado como la quinua.

Esta investigación es novedosa en el sentido de que, de acuerdo a los trabajos revisados, es uno de los primeros intentos enfocados en la clasificación de semillas de quinua usando técnicas de procesamiento digital de imágenes e inteligencia computacional, evaluando algoritmos de clasificación para poder agrupar de manera automática semillas de quinua de tres tipos (Sacaca o amarilla, Pasankalla o roja y Salcedo). Inicialmente se capturaron 360 imágenes de semilla de quinua con fondo blanco y negro, después con la ayuda de un ingeniero agrónomo se caracterizaron obteniendo 199 de la quinua Sacaca, 115 de la quinua Pasankalla y 127 de la quinua Salcedo; según su validación entre buenas y malas se tomaron como positivas y negativas para el entrenamiento de los clasificadores.

Inicialmente se compararon filtros para corregir y cuantificar el color en las imágenes permitiendo resaltar las diferencias entre las características de cada tipo de Semilla de Quinua, los filtros comparados fueron Medianblur, Gaussianblur y blur. Después de corregir el color de las imágenes se realizó la segmentación utilizando el método de Otsu con el fin obtener una mejor extracción de características para el entrenamiento de los clasificadores.

Para lograr la clasificación automática de las semillas de quinua se implementó el clasificador SVM (Support Vector Machines) obteniendo una precisión de 66%.

2 Trabajos relacionados

En la investigación [5] se utilizaron técnicas para la identificación de nueve variedades de semillas de trigo iraní por análisis textural con procesamiento de imágenes, varios grupos de características texturales de imágenes de semillas fueron examinados para evaluar su eficacia en la identificación de semillas. En conjunto, 1080 imágenes en escala de grises de semillas de trigo a granel (120 imágenes de cada variedad) fueron adquiridos en una condición estable de iluminación (luz fluorescente anillo). Se obtuvieron un total de 131 características texturales, nivel gris, gray level cooccurrence matrix (GLCM), gray level run length matrix (GLRM), local binary patterns (LBP), local similarity patterns (LSP) y matrices local similarity numbers (LSN). Se empleó el método de supuesta discriminación paso a paso para seleccionar y clasificar las características texturales más significativas de cada matriz individualmente, así como las características de todas las matrices simultáneamente. El Clasificador lineal discriminante analysis (LDA) fue empleada para la clasificación utilizando características seleccionadas. La exactitud de la clasificación promedio de 98.15% se obtuvo cuando se utilizaron en el clasificador 50 de las mejores características. Los resultados confirmaron que características LSP, LSN y LBP tuvieron una influencia significativa en la mejora de la exactitud de la clasificación en comparación con estudios previos.

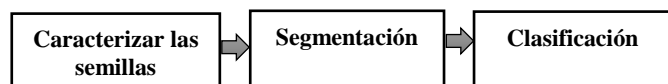
En el estudio [6] se empleó imágenes hiperspectrales para desarrollar métodos de clasificación de semillas de maíz con el objetivo de aumentar el rendimiento de las cosechas. Se evaluaron 2000 imágenes de semillas correspondientes a 4 variedades de maíz, las imágenes de reflectancia hiperspectral se adquirieron entre 400nm y 1000 nm.

Los modelos de clasificación basados en las características espectrales medias de las semillas se desarrollaron usando least squares support vector machine (LSSVM). La actualización de modelos utilizando datos vectoriales de apoyo incrementales también se aplica para actualizar el modelo LSSVM en línea y asegurar la identificación precisa de semillas de maíz de diferentes años.

La precisión de clasificación del modelo LSSVM combinada con la actualización del modelo fue 94,4%, un 10,3% mayor que la de otros modelos no actualizados.

3 Metodología

El método propuesto para la clasificación de los tipos de semillas de quinua consta de 5 etapas.



a. Caracterizar las semillas

Se contó con la ayuda de un ingeniero agrónomo experto para determinar las características más importantes de los tres tipos de semillas de quinua.

1. Quinua Sacaca o Amarilla

La validación según el experto es la siguiente:

Estado bueno: La semilla muestra un color en el epispermo blanco amarillento a amarillo anaranjado; un diámetro de 1,6 a 2,2 mm y su forma granular está bien definida, (Fig. 1-a).

Estado malo: la semilla muestra oscurecimiento en zonas del epispermo en algunos casos con presencia de manchas, además de una pérdida en el diámetro y su forma granular es indefinida, (Fig. 1-b).

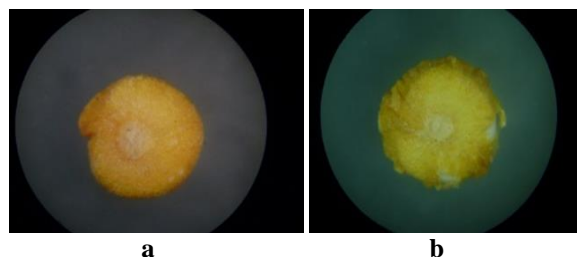


Fig. 1: Semilla de quinua Sacaca en estado bueno (a), estado malo (b)

2. Quinua Pasankalla o Roja

La validación según el experto es la siguiente:

Estado bueno: La semilla muestra un color en el epispermo rojo vino, su pericarpio es plomo claro y un diámetro de 1,9 - 2,0 mm, su forma granular está bien definida., (Fig. 2-a).

Estado malo: la semilla muestra oscurecimiento en zonas del epispermo en algunos casos con presencia de manchas, además de una pérdida en su forma granular así también de parásitos debido a que su sabor es dulce, (Fig. 2-b).

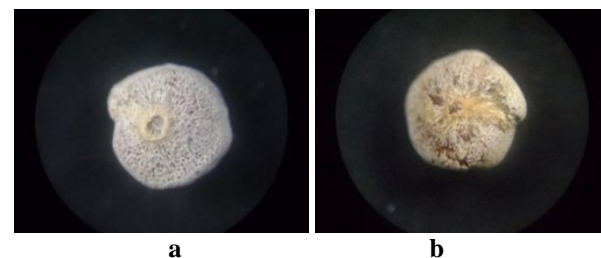


Fig. 2: Semilla de quinua Pasankalla en estado bueno (a), estado malo (b)

3. Quinua Salcedo

La validación según el experto es la siguiente:

Estado bueno: La semilla muestra un color en el epispermo blanco, su pericarpio es verdoso a amarillento y un diámetro de 1,5 - 2,2 mm, (Fig. 3-a).

Estado malo: la semilla muestra oscurecimiento en zonas del epispermo en algunos casos con presencia de manchas y protuberancias debido a hongos o parásitos que infectan su interior, además de una reducción en el diámetro y su forma granular es indefinida (Fig. 3-b).

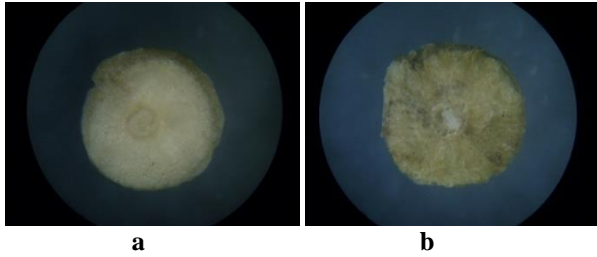


Fig. 3: Semilla de quinua Salcedo en estado bueno (a), estado malo (b)

b. Segmentación



Fig. 4: Método de segmentación

Gaussianblur

Se han comparado filtros para corregir y cuantificar el color de las imágenes de forma que permitieran resaltar las diferencias entre las características de cada tipo de semilla de quinua.

Se evaluaron los filtros Median Blur, Gaussianblur y blur; de estos el que obtuvo mejores resultados fue el filtro Gaussianblur [7].

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

Donde:

- x : Distancia desde el origen en el eje horizontal
- y : Distancia desde el origen en el eje vertical
- σ : Desviación estándar

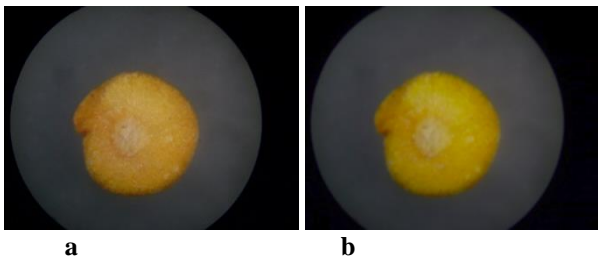


Fig. 5: Imagen de semilla de quinua (a), imagen con filtro Gaussian blur (b)

Segmentación

Para diferenciar la semilla de quinua del fondo oscuro de la imagen ($f(x,y)$) se utilizó la binarización, empleando la Ecuación (2).

$$g(x, y) = \begin{cases} 1 & \text{si } f(x, y) > T \\ 0 & \text{si } f(x, y) \leq T \end{cases} \quad (2)$$

Para determinar el valor del umbral (T) se utilizó el método de OTSU.

$$n(t) = \frac{\sigma_b^2}{\sigma_d^2}$$

Sea σ_b^2 e σ_d^2 la varianza entre clases y el histograma [8].

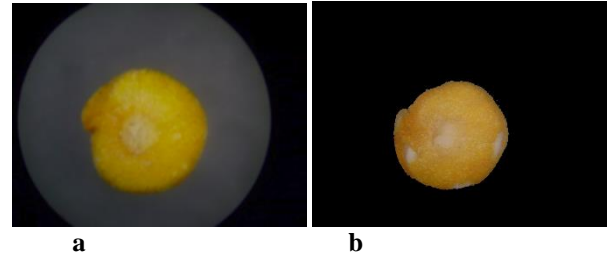


Fig. 6: Imagen de semilla de quinua (a), imagen segmentada (b)

c. Extracción de características

Luego de someter la imagen a la segmentación esta se convierte a escala de grises y se redimensiona para liberar conflictos con la memoria interna, después se obtiene las características mediante una transformación que convierte el espacio de característica Matriz (W,H) en un matriz de una sola fila (1, W x H) para ser ingresada en cada fila de la matriz de entrenamiento cuya conversión genera un valor numérico relacionado a los canales en lo se encuentra los pixeles en el espacio característico en este caso RGB y que sirve como peso en los algoritmos por los cuales entrena los clasificadores de ese modo permiten hallar diferencias entre los objetos a clasificar en este caso la semilla de quinua.

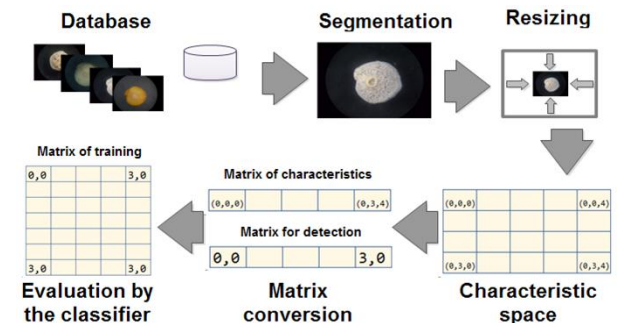


Fig. 7: Matriz de entrenamiento

La matriz de entrenamiento y la de características coincide en sus dimensiones siendo en el ancho y el alto según corresponde a la multiplicación de las dimensiones del espacio característico de la imagen redimensionada y el número de imágenes de estas misma.

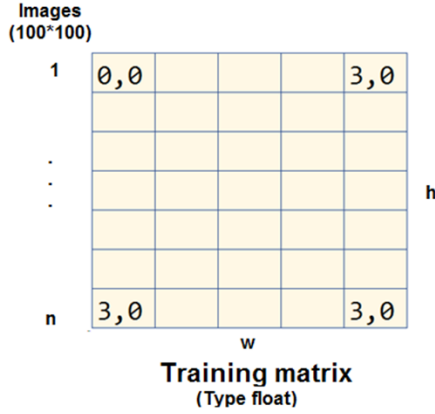


Fig. 8: Dimensiones de la matriz

Se etiquetaron las imágenes, 1 (Positivas o Buenas) y -1 (Negativas o Malas), según el tipo de semilla de quinua en su validación comenzando por la Quinua Amarilla o Sacaca, la Quinua Roja o Pasankalla, la Quinua INIA Salcedo y continuaría así por cada tipo solo que para el objetivo de esta investigación se tomo estas tres como muestra.

Después se asigna estas etiquetas a la fila que hace referencia a la imagen ingresada a la matriz de entrenamiento.

white quinua - INIA Salcedo - good	3	0,0				3,0
red quinua - Pasankalla - good	2					
yellow quinua - Sacaca - good	1					
Does not belong to any type	0					
yellow quinua - Sacaca - bad	-1					
red quinua - Pasankalla - bad	-2					
white quinua - INIA Salcedo - bad	-3	3,0				3,0
Labels (matrix 1 * n)		Training matrix (Type float)				

Fig. 9: Etiquetado de la matriz

Para la detección de características la imagen seleccionada debe pasar por un preprocesamiento que genera una imagen segmentada, luego esta se convierte a escala de grises para ser redimensionada y la matriz de espacios característico se extrae en una matriz de una sola fila para ser introducida a la matriz de características y convertida en una matriz de detección que al final es evaluado por el algoritmo de que clasificador ya entrenado para su respuesta que es la etiqueta a la cual pertenece.

d. Clasificación

Para la clasificación automática de los tipos de semillas de quinua se utilizó Support Vector Machine (SVM)[9].

La clasificación se realizó en siete clases $y_i \in \{-3, -2, -1, 0, 1, 2, 3\}$, las imágenes evaluadas tienen un tamaño de 100x100 pixeles entonces $x_i \in \mathbb{R}^{10000}$.

Solución:

$$\omega = \sum_{i=1}^7 \alpha_i y_i x_i^{sp} \quad (4)$$

Tipo semilla:

$$f(x_{new}) = \text{sign}\left(\sum_{i=1}^7 \alpha_i y_i \langle x_i^{sp}, x_{new} \rangle + b\right) \quad (5)$$

4 Experiments And Results

Materiales

Las semillas de quinua fueron obtenidas de los andes peruanos.

Las imágenes fueron digitalizadas utilizando la cámara SONY Cyber-Shot DSC-S750 sobre el Estereoscopio 40x, en un entorno especialmente diseñado para mitigar cualquier sombra que pudiera afectar el muestreo [10].

El entorno se crea a partir de una base donde reposa la semilla, la cual esta alumbrada de un modo que se evite el reflejo de sombras a su vez la imagen es aumentada por el estereoscopio para evitar que la cámara distorsione los megapíxeles debido a que los tipos de semilla de quinua son muy pequeños y de esta forma sea más fácil el proceso de muestreo.

Para validar el tamaño se utilizó un calibrador llamado Vernier o Pie de Rey ya que por ser una herramienta destinado para la medición de pequeñas longitudes, espesores, profundidades y diámetros interiores.

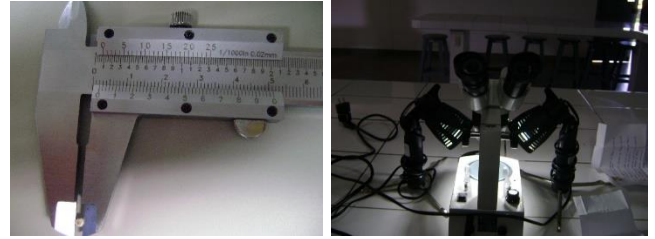


Fig. 10: Herramientas y materiales

Las imágenes tomadas poseen una medida de 2048 x 1536 pixeles aproximadamente. Se capturaron imágenes de semilla de quinua con fondo blanco y negro; después con la ayuda de un ingeniero agrónomo y según el estándar establecido por el Instituto Nacional de Innovación Agraria (INIA Perú) se caracterizaron obteniendo 199 de la quinua Sacaca, 115 de la quinua Pasankalla y 127 de la quinua Salcedo.

Resultados

This section will present the results associated to the classification of quinoa seed types. Specifically, it will be presented information related to the amount of correctly classified and incorrectly classified images, and the global accuracy. In addition, it will be use the evaluation metrics precision, recall and F1 [11] obtained for each class, in this case, for each type of seed quinoa. In all case it was used a 10-fold cross validation approach.

Figure 9 presents the definition for precision, recall and F1 in terms of true-positive, false-positive, false-negative, and true-negative results. The exact meaning of such results is visualized in the confusion matrix (table 1).

$$precision = \frac{\#tp}{\#tp + \#fp} \quad (6)$$

Table 1:
Confusión matrix

		Prediction	
		Positive	Negative
Actual	Positive	TP	FN
	Negative	FP	TN

Procesamiento

Se han comparado filtros para corregir y cuantificar el color de las imágenes de forma que permitieran resaltar las diferencias entre las características de cada tipo de semilla de quinua.

Table 2:
Comparación de filtros.

Filtros	MOS de diferencia aceptables		
	Imagen 1	Imagen 2	Imagen 3
Medianblur	1.8	2.5	3.6
Gaussianblur	1.4	2.2	3.4
Blur	2.7	3.2	3.8

Según la tabla 2, se comparó los filtros de corrección, concluyendo que Gaussianblur es el que posee la menor diferencia, siendo el más óptimo para este propósito.

Table 3:
Results using SVM classification

Clasificación con el algoritmo SVM

Type	Images	TP	TN	Precision %
Quinua Amarilla	100	48	11	59
Quinua Roja	112	68	13	72
Quinua Salcedo	86	34	23	66
Total aciertos	298	150	47	66

Utilizando la clasificación con SVM se obtiene una precisión promedio de 66% para la clasificación de los tipos de semilla de quinua.

5 Conclusiones

El protocolo de la captura de las imágenes de semillas de quinua permitió digitalizarlas con mejor calidad para el procesamiento, siendo el uso de la iluminación, microscopio y cámara fotográfica las mejores herramientas.

Después de la captura de las imágenes es necesario corregir y cuantificar el color para resaltar las diferencias entre las características de cada tipo de semilla de quinua, el filtro que tiene mejores resultados para esta tarea es Gaussianblur.

se comparó los filtros de corrección, concluyendo que Gaussianblur es el que posee la menor diferencia, siendo el más óptimo para este propósito.

Utilizando la clasificación con SVM se obtiene una precisión promedio de 66% para la clasificación de los tipos de semilla de quinua.

Agradecimientos

Agradecimiento por su apoyo en la caracterización de las imágenes de quinua al Ingeniero agrónomo Jorge Leyva Piedra del Parque Tecnológico Científico de la Universidad Señor de Sipán.

Referencias

- Carrillo, E.; Aristizábal Peñaloza, "Artificial vision to assure coffee-Excelso beans quality," 2009.
- Daza, L. Sánchez G., and J. Suárez C., "Selección De Características Orientada a Sistemas De Reconocimiento De Granos Maduros De Café," *En Rev. Sci. Tech. Año XIII*, vol. 35, no. 35, pp. 139–144, 2007.
- Goutte, C.; Gaussiel, "A probabilistic interpretation of precision, recall and F-score, with implication for evaluation," *Proc. Eur. Conf. Inf. Retr.*, pp. 345–359, 2005.
- Hsu, C. Chang, and C. Lin, "A Practical Guide to Support Vector Classification," vol. 1, no. 1, pp. 1–16, 2016.
- Huang, J. Tang, B. Yang, and Q. Zhu, "Classification of maize seeds of different years based on hyperspectral imaging and model updating," *Comput. Electron. Agric.*, vol. 122, pp. 139–145, 2016.
- Jacobsen, A. Mujica, and C. R. Jensen, "The Resistance of Quinoa (*Chenopodium quinoa* Willd.) to Adverse Abiotic Factors," *Food Rev. Int.*, vol. 19, no. 1–2, pp. 99–109, 2003.
- Katty, T. "Demanda de la Quinua (*Chenopodium quinoa* Willdenow) a nivel industrial," Universidad Agraria la Molina - Perú, 2003.

- Panesar, J.P.; Rana, and S. Singh, "Structural, thermal and rheological properties of starches isolated from Indian quinoa varieties," *Int. J. Biol. Macromol.*, vol. 102, pp. 315–322, 2017.
- Pourreza, H. Pourreza, M. H. Abbaspour-Fard, and H. Sadrnia, "Identification of nine Iranian wheat seed varieties by textural analysis with image processing," *Comput. Electron. Agric.*, vol. 83, pp. 102–108, 2012.
- Ro and R. T. A. Hummel, "Deblurring Gaussian Blur," vol. 80, pp. 66–80, 1987.
- Smith, D. B. Reid, C. Environment, L. Palo, P. Alto, and P. L. Smith, "A Threshold Selection Method from Gray-Level Histograms," vol. C, no. 1, pp. 62–66, 1979.