

# Clasificación automática de arándanos con daños fúngicos

Gabriel Leiva Valenzuela<sup>(1)</sup>, Germán Mondragón Gonzalez<sup>(2)</sup>, Cristóbal Arrieta Pellegrin<sup>(3)</sup>, Domingo Mery<sup>(2)</sup>

(1)Departamento de Ingeniería Química y Bioprocesos.

(2)Departamento de Ciencias de la Computación.

(3)Departamento de Ingeniería Química.

Pontificia Universidad Católica de Chile

Santiago, Chile

dmery@ing.puc.cl

**Abstract—** La producción de arándano sudamericano ha aumentando sobre un 40% en los últimos años movido por la demanda internacional. Sin embargo, la eficiencia de las exportaciones disminuyen por la selección manual de calidad motivando a la industria a buscar alternativas más confiables. En este trabajo implementamos algoritmos de inspección de imágenes fotográficas para segregar arándanos que presentaron pudrición de unidades en buen estado. Obtuvimos un 98% de rendimiento para el reconocimiento de daños fúngicos usando clasificador LDA, validando las mediciones por el método de “10-fold cross validation” con un 95% de confianza [92-100]. Adicionalmente se clasificaron arándanos dañados en general de arándanos buenos obteniendo valores de 84% de rendimiento. La implementación de estos sistemas en arándanos fue compleja debido a su morfología, color, tamaño y tipo de daños producidos luego de la cosecha, abriendo una importante veta de investigación que permitirá mejorar la competitividad del sector frutícola cuando sean implementados en la línea de producción.

**Keywords-** arándanos; pudrición; poscosecha; clasificación; supervisión; fotografía.

## I. INTRODUCCION

En las últimas décadas las exigencias de calidad de productos frutícolas han aumentado junto con las tecnologías de selección, preservación, y transporte de los *commodities*. Esto ha creado la necesidad de incluir automatización y mejorar el control eficiente de su calidad. En manzanas, por ejemplo, se han implementado sistemas de selección automáticos en industrias exportadoras a través de procesos de supervisión en línea de imágenes a color e infrarrojo cercano, para clasificar productos según calibre, color y defectos obteniendo resultados más confiables y a mayor velocidad que la selección manual. En cerezas los clasificadores automáticos incluyen sólo el color promedio y la forma.

Chile ha llegado a producir 42 mil toneladas anuales de arándanos (*Vaccinium corymbosum*) en el periodo 2008-2009, creciendo un 66% desde el periodo anterior, enfocando su producción para el mercado exportador, fundamentalmente EE.UU. [1]. Este aumento en las exportaciones demanda mayores esfuerzos en asegurar la calidad del producto.

Chile y Argentina son los mayores productores de arándano fresco del hemisferio sur creciendo a tasas sobre el

40% anual debido a la demanda de los EE.UU., Canadá, y Europa en su periodo invernal-primaveral. Estos mercados también son importantes productores, sin embargo, no pueden auto-abastecerse fuera de temporada debido a la corta vida útil del fruto.

El arándano es un frutal menor de la familia de las mirtáceas del tipo *berries*. El fruto es esférico, tiene un diámetro promedio de 1.5 cm, su epidermis tiene color azul oscuro pudiendo estar recubierta con ceras naturales. El fruto del arándano es comestible, tiene un sabor agridulce característico y tiene componentes antioxidantes de la familia de las antocianinas las que le confieren su color azul característico.

Los arándanos son propensos a dañarse luego de la cosecha. Sus características intrínsecas como la epidermis delgada, su diámetro pequeño, su alto contenido de acidez, y su textura, así como los tratamientos de selección lo hacen susceptibles a daños por mohos.



**Fig. 1.** Pudrición en arándanos.

La pudrición es un daño de tejidos internos y externos producido por microorganismos (Fig. 1). Tanto hongos, levaduras o bacterias son inoculados desde la superficie de la epidermis hacia el interior de los arándanos por lesiones que se producen en la zona pedicelar cuando los frutos son cosechados. El inóculo puede desarrollarse si las condiciones ambientales le son favorables por ejemplo deficiencias en el enfriado rápido sumado a altas temperaturas. Los principales microorganismos que producen la pudrición en arándanos son mohos como *Botrytis cinerea*, *Rhizopus stolonifer* o *Colletotrichum spp*. En la fase mayor de desarrollo de la pudrición, se observan

sectores con micelio fúngico visible en la zona pedicelar del fruto. Implementar planes de segregación de unidades con pudrición en fases tempranas permite evitar la propagación del deterioro durante el transporte evitando importantes pérdidas económicas.

Otros daños visuales de poscosecha como deshidratación y daño mecánico, junto con otros no visuales como ablandamiento están relacionados con la pudrición, pudiendo encontrarse en un mismo fruto varios daños. La pudrición puede ser precursor de los otros daños.

En la actualidad, la selección de calidad en arándanos para exportación es realizada por personas que toman en consideración el color, tamaño, forma, y daños poscosecha. Esta práctica es subjetiva, lenta y es incapaz de segregar correctamente la totalidad de los frutos.

Abbott [2] señala que la variabilidad entre frutas requiere la separación por clases de frutos. Las frutas de un mismo huerto y variedad tienen una gran variabilidad fenotípica, diferenciándose cada unidad del promedio. Por esta razón, los planes de muestreo son insuficientes, necesitando procesos de separación por clases que incluyan la totalidad de las unidades que aseguren la uniformidad requerida por el mercado. Esta clasificación requiere alta velocidad, sensibilidad, confiabilidad que pueden ser efectuados por técnicas no destructivas basadas en la adquisición de imágenes.

La clasificación de arándanos según presencia de pudrición presenta los siguientes desafíos intrínsecos del fruto:

- color oscuro con cantidades variables de cera blanca natural (pruina);
- tamaño pequeño del fruto (1,4 cm. de diámetro promedio);
- diferencias en el tamaño del sector de pudrición;
- visualmente confusión en algunos sectores entre la pudrición y el sector pedicelar (unión del fruto con la planta).

Al implementar mecanismos no destructivos de clasificación de en plantas de procesamiento para la exportación ocurrirían ventajas en los siguientes aspectos:

- 1) Se logrará exportar frutos más homogéneos;
- 2) Se segregarán arándanos por estándares de calidad según mercados de exportación;
- 3) Se aumentará el tiempo de almacenamiento de las clases de mejor calidad;
- 4) Se disminuirán las pérdidas y;
- 5) Se disminuirán los costos de mano de obra.

Estos cinco aspectos en conjunto mejorarían la competitividad y el retorno económico de la inversión necesaria para su implementación.

Actualmente, para el control de calidad, la selección de arándanos se realiza de forma manual mediante personal temporal con baja calificación, lo que se traduce en una alta

tasa de fallas en la selección, disminuyendo la calidad del producto final.

Una buena clasificación en etapa de poscosecha es fundamental para la calidad de este producto de exportación. Se persigue de esta manera que el arándano exportado arribe al lugar de destino con los estándares de calidad requeridos por el mercado internacional. Con una mala clasificación aumenta el riesgo de rechazos de lotes de producción por contaminación y propagación del daño.

En esta investigación, pensamos que el uso de algoritmos de clasificación automática podría resolver el problema del control de calidad de arándanos, ya que existen antecedentes de frutas clasificados mediante reconocimiento de patrones que permiten predecir un buen resultado (Tabla 1).

**Tabla 1** Control de calidad automático usando imágenes ópticas

Fruto	Aplicación
Manzanas	Segmentación de defectos, Color de fondo, tamaño
Cerezas	Tamaño, color promedio
Granos	Clasificación de calidad
Tomates	Madurez por color
Pistachos	Detección de partidura prematura
Espárragos	Inspección de defectos
Cerezas	Análisis de forma

Fuente: Adaptado de Du y Sun [7]

Por esta razón utilizamos la metodología estándar de reconocimiento de patrones para realizar la clasificación automática de arándanos con daños fúngico. Los pasos de esta metodología son:

- 1) Pre-procesamiento de imágenes fotográficas tomadas en condiciones estándar [3, 4],
- 2) Extracción de características geométricas y cromáticas,
- 3) Selección de características [5],
- 4) Diseño de clasificadores [6] y
- 5) Evaluación del desempeño de distintos clasificadores.

Con esto clasificamos arándanos en dos clases: en buen estado visual y en mal estado visual. Se obtuvieron rendimientos de un 98% considerando como mal estado aquellas unidades que presentaron pudrición fúngica (por hongos).

Se estima que realizar una clasificación automatizada de este defecto disminuiría la tasa de fallas en una selección posterior que considere defectos no visuales.

El artículo presentado está organizado de la siguiente manera: la sección II presenta los materiales y métodos empleados, la sección III muestra los resultados experimentales y la sección IV las conclusiones obtenidas en esta investigación.

## II. MATERIALES Y MÉTODOS PROPUESTOS

### A. Materiales

Para la experimentación se utilizaron arándanos (*Vaccinium corymbosum* var. Highbush) obtenidos de huertos Chilenos. Los arándanos fueron mantenidos por dos meses en cámara de frío a 0° C para producir distintos daños por almacenaje. La toma de fotografía fue realizada en condiciones ideales mediante colorímetro DVS-lab (de la empresa Divisol [www.divisol.cl](http://www.divisol.cl)). Los datos fueron analizados en Matlab (Mathworks corp.).

### B. Adquisición de imágenes

Los arándanos fueron retirados de cámara de conservación y clasificados entre arándanos buenos y malos. Se consideró como arándanos malos aquellas frutas que presentaron daño por pudrición, deshidratación y daño mecánico. Esto fue realizado por inspección visual por personal calificado. A cada grupo dividido se le tomó una fotografía por la zona del cáliz y otra por la zona pedicelar.

### C. Segmentación de imágenes

La segmentación consistió en generar una imagen de alto contraste monocromática de a partir de una imagen RGB [4]. Cada imagen segmentada fue dividida generando una imagen por arándano mediante el cálculo de una región rectangular a partir de los centros de masa de sus imágenes.

### D. Extracción de características

Se extrajeron 1643 características geométricas y cromáticas en los espacios de color R, G, B, Hue, Saturation, Value, L\*, a\*, b\* y tonos de gris [8]. Estas características incluyen texturas, Gabor, transformadas de Fourier y discreta de cosenos y contrastes.

### E. Selección de características

Se utilizaron diversas estrategias para la selección de características tales como SFS [6] (con función objetivo del tipo discriminante de Fisher y también con funciones objetivo obtenidas de los desempeños de los clasificadores red neuronal, QDA, LDA y KNN), FOS [9], LSEF [10], Rankfs [11], con distintas funciones objetivos de minimización [12].

### F. Clasificación

Las características seleccionadas por los siete métodos fueron clasificadas por diversos clasificadores tales como red neuronal [13], LDA, Mahalanobis, QDA, SVM, vecino más cercano (KNN) con 10 y 30 vecinos, y distancia mínima [5].

### G. Evaluación de desempeño

Para evaluar el desempeño de los clasificadores se utilizó la técnica de “10-cross validation” con un 95% de confianza.

## III. EXPERIMENTOS Y RESULTADOS

Los experimentos realizados en esta investigación se dividieron en dos partes:

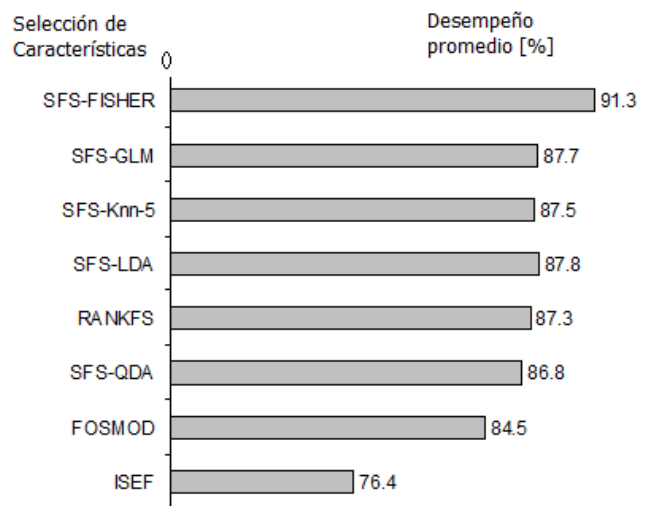
- 1) selección de estrategia y del número de características, y
- 2) clasificación-validación estadística del método

Los métodos se probaron en la clasificación de:

- a) arándanos con pudrición y
- b) arándanos con pudrición más deshidratación y daño mecánico.

La Fig. 2. muestra los resultados que permitieron determinar cuál de las estrategias de selección de características (sección II.E) sería la utilizada en la obtención de resultados posteriores. En esta figura se muestra los rendimientos promediados de 22 clasificadores para imágenes adquiridas de cáliz utilizando las mejores 8 características seleccionadas en cada estrategia.

Al evaluar el rendimiento para la clasificación de daños fúngicos de arándanos mediante distintas estrategias y distintos clasificadores, observamos en general un mejor rendimiento con la estrategia SFS-Fisher tanto para cáliz como las imágenes de pedúnculo de arándanos. Esta estrategia consistió en elegir las 8 mejores características con SFS desde 1643 características totales.



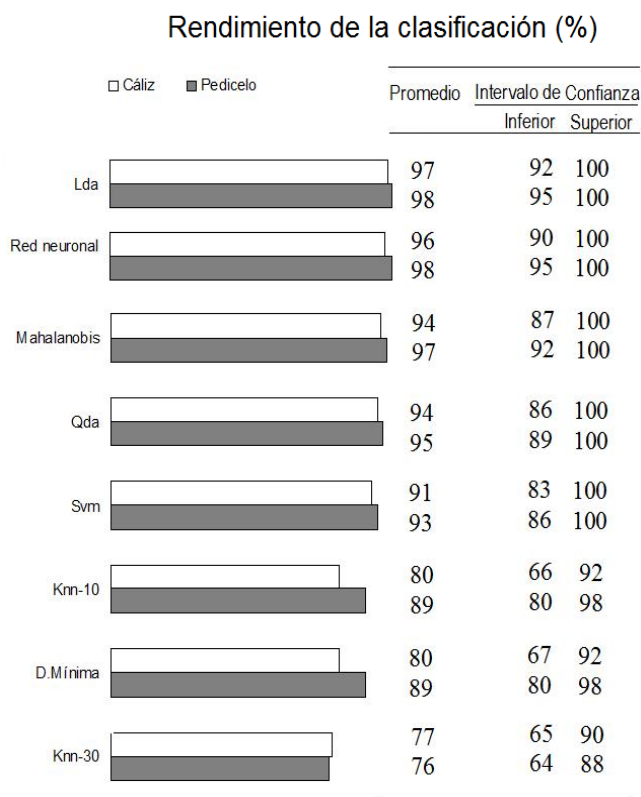
**Fig. 2.** Rendimiento promedio para distintas estrategias de selección de características.

**Tabla 2** características escogidas para la selección de arándanos con pudrición en zona pedicelar y de cáliz.

Pedicelo	Cáliz
Tx 2,d1(mean) -Hue	Tx 2,d1(mean) -Hue
Tx 1,d1(mean) -Hue	Eccentricity
Hu-moment 3	Danielsson fac
Gabor(1,3) -Gray	Gabor-J -Hue
Tx 2,d2(mean) -Green	Fourier_3 -Sat
DCT_4 -Gray	Hu-moment 3
Fourier-descr 10	Tx 8,d1(mean) -Red
Int-Hu-moment 3 -Sat	Tx 14,d1(range) -Sat

En la Tabla se muestran las 8 características seleccionadas tanto para la clasificación automática de Pedicelo como Caliz. Se observa la selección tanto de características geométricas (momentos de Hu, excentricidad, etc.) como cromáticas (Texturas de Haralick, Gabor, Fourier, etc ) [8].

La Fig. 4 muestra los resultados de la clasificación de arándanos buenos y malos usando distintos clasificadores. Se evaluaron los intervalos de confianza con un 95% de confiabilidad mediante la técnica de validación cruzada. Observamos que el mejor clasificador fue LDA seguido de la red neuronal utilizada con rendimientos superiores al 96% tanto para imágenes de zona pedicelar como de cáliz.



**Fig. 4.** Resultados obtenidos en la clasificación.

Al clasificar los arándanos entre buenos y malos considerando la totalidad de los defectos incluidas pudriciones, deshidrataciones y daños mecánicos, se obtiene para imágenes de pedicelo un 84,7% de rendimiento con un intervalo de confianza entre 76,7% y 92,6% utilizando una estrategia de extracción con SFS-QDA y un clasificador SVM con un 95% de confianza.

La misma evaluación para imágenes de cáliz muestra un rendimiento de 89,9 % con intervalos de confianza que van desde 83,7% a 96,2% utilizando la extracción de característica SFS-Fisher con un clasificador LDA.

Estos datos por primera vez muestran un mayor rendimiento en la zona de cáliz que en el pedicelo, pudiendo deberse a que la presencia de otros defectos se acentúa en dicha zona. Sin embargo los resultados no alcanzan el 95% de rendimiento necesario para que la selección sea eficiente.

#### IV. CONCLUSIONES

Los arándanos son frutos que tienen los siguientes atributos que complican la detección de defectos mediante visión por computador: bajo calibre, alta variación de características fenotípicas intra cultivar, color oscuro, cantidades fluctuantes de ceras de color blanco. Los principales daños poscosecha como la pudrición por mohos, y otros de menor incidencia como deshidratación, el ablandamiento y los daños mecánicos son difíciles de clasificar mediante técnicas convencionales de selección por medio de personas, lo que ha traído consecuencias económicas por rechazo de lotes por no conformidad. En este trabajo se buscó como primera aproximación la clasificación de arándanos con daño por hongos para ser separado en línea de proceso.

Las características extraídas fueron 1643, sin embargo sólo con una porción reducida de ellas es posible realizar una clasificación efectiva. La etapa de selección de características muestra que SFS-fisher entrega los mayores rendimientos al seleccionar 8 características. Esta información es relevante al momento de programar un algoritmo de clasificación para ser implementado en tiempo real en línea de procesos al disminuir los tiempos de extracción desde

Los mejores clasificadores tanto para las imágenes adquiridas de zona pedicelar y de cáliz fueron LDA con un 98% y un 97% de rendimiento promedio respectivamente. Los intervalos de confianza fueron calculados mediante validación cruzada con un 95% de confianza fluctuando entre un 92% de rendimiento como límite inferior y un 100% de clasificación como límite superior. Es interesante notar que la simplicidad del clasificador lineal entregó

mejores resultados que otros de mayor complejidad como el clasificador de red neuronal o SVM.

En adición a la clasificación de arándanos con daño fúngico, se evaluaron daños menores como deshidratación o daño mecánico encontrando

Como perspectiva futura, se debe mejorar técnicas de clasificación los daños menos evidentes como deshidratación y daño mecánico los que presentaron valores de rendimiento menores al 90%. Adicionalmente, se debe profundizar en la clasificación de unidades que visualmente se ven bien, pero presentan el defecto no visible llamado ablandamiento. Esto podría ser solucionado con técnicas no visibles como infrarrojo cercano, ultrasonido, rayos X o Resonancia Nuclear magnética las que al ser implementadas podría constituir un aporte para el desarrollo de la industria frutícola del arándano.

#### REFERENCIAS

- [1] ODEPA, (2009), "Situación de los mercados de exportación de tres frutas en expansión: paltas, arándanos y cerezas".
- [2] Abbott, J.A.,(1999). Quality measurement of fruits and vegetables. *Postharvest Biology and Technology* 15: 207-225.
- [3] Golnabi y Asadpour A.H. and A. Asadpourb, (2007), Design and application of industrial machine vision systems. *Robotics and Computer-Integrated Manufacturing* 23: 630-637.
- [4] Mery, D. and Pedreschi, F. (2005). Segmentation of color food images using a robust alg orithm. *Journal of Food Engineering*. 66: 353-360.
- [5] Duda, R., Hart, P., Stork, D.(2006) *Pattern Classification*, New York, John Wiley & Sons, Inc., 2001. Bishop, C.: *Pattern Recognition and Machine Learning*, Springer, 2006
- [6] Jain, A., Duin, R., Mao, J. (2000), "Statistical Pattern Recognition: A Review", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 1.
- [7] Du, C.J. and DW. Sun, (2004). Recent developments in the applications of image processing techniques for food quality evaluation. *Trends in Food Science & Technology* 15: 230-249.
- [8] Mery, D., Soto, A. (2008): *Features: The more the better*. The 7th WSEAS International Conference on Signal Processing, Computational Geometry and Artificial Vision (ISCGAV-2008), Rodos Island, Greece, August 20-22.
- [9] Wei, H.-L. and Billings, S. Feature Subset Selection and Ranking for Data Dimensionality Reduction *Pattern Analysis and Machine*.
- [10] Mao, K., (2005) Identifying critical variables of principal components for unsupervised feature selection Systems, Man, and Cybernetics, Part B: Cybernetics, *IEEE Transactions on*. 35, 339-344
- [11] Liu, H., Motoda H. (1998) *Feature Selection for Knowledge Discovery and Data Mining*, Kluwer Academic Publishers
- [12] Duda, R., Hart, P., Stork, D.(2006) *Pattern Classification*, New York, John Wiley & Sons, Inc., 2001. Bishop, C.: *Pattern Recognition and Machine Learning*, Springer, 2006
- [13] Bishop, C.(2006). *Pattern Recognition and Machine Learning*, Springer, 2006