



Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

**ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA AGRONÓMICA Y
BIOCIENCIAS**

***NEKAZARITZAKO INGENIARITZAKO ETA BIOZIENTZIETAKO GOI MAILAKO
ESKOLA TEKNIKO***

CLASIFICACIÓN DE PAPA DE ACUERDO CON SU APTITUD DE TRASFORMACIÓN
INDUSTRIAL FRITO Y COCIDO USANDO IMÁGENES HIPERESPECTRALES

presentado por

GIL ALFONSO OLIVA LOBO

aurkeztua

MÁSTER UNIVERSITARIO EN TECNOLOGÍA Y SOSTENIBILIDAD EN LA INDUSTRIA ALIMENTARIA
UNIBERTSITATE MASTERRA ELIKAGAIEN INDUSTRIAKO TEKNOLOGIAN ETA JASANGARRITASUNEAN

Junio, 2020 / 2020ko, ekaina

RESUMEN

Desde un punto de vista logístico, económico y teniendo en cuenta que la papa es un cultivo de gran importancia para la economía global, surge la necesidad de caracterizar sus propiedades y aptitudes a fin de cumplir con los estándares de calidad y seguridad establecidos. Las imágenes hiperespectrales (HSI) para la clasificación de alimentos es una de las técnicas más utilizadas en la actualidad, ya que permite la medición simultánea de datos físicos y espectrales sin destruir la muestra. Por ello el objetivo de este estudio fue evaluar la funcionalidad de imágenes hiperespectrales para clasificar de distintas variedades de papa en función de su aptitud de fritura y cocción. Se utilizaron 72 tubérculos de 18 variedades diferentes y se tomaron imágenes en el rango 900-1700 nm. Mediante un análisis discriminante de mínimos cuadrados parciales (PLS-DA) se desarrolló un modelo capaz de clasificar correctamente el 66,44% de papas para la aptitud cocción y el 64,38% para la aptitud fritura.

PALABRAS CLAVE

Imágenes hiperespectrales, Papa (*Solanum tuberosum* L.), análisis PLS-DA, Fritura, Cocción.

ABSTRACT

From a logistical, economic point of view and keeping in mind that potatoes are a crop of great importance both for food security and for the global economy, there is an increasing need to characterize their properties and aptitudes in order to comply with the quality and safety standards established for this crop. Hyperspectral imaging (HSI) for food classification is currently one of the most used techniques, since it allows the simultaneous measurement of physical data and spectral data without destroying the sample. Therefore, the objective of this study was to evaluate the functionality of hyperspectral images to classify different potato varieties based on their aptness to fry and cook. 72 tubercles of 18 different varieties were measured in a hyperspectral system in the 900-1700 nm range and through a partial least square's discriminant analysis (PLS-DA) a model was developed able to classify 66.44% of potatoes for cooking aptitude and 64.38% for frying aptitude.

KEY WORDS

Hyperspectral images, Potato (*Solanum tuberosum* L.), PLS-DA analysis, Frying, Cooking.

ÍNDICE GENERAL

1.	INTRODUCCIÓN Y OBJETIVO	6
2.	REVISIÓN BIBLIOGRÁFICA	8
2.1	Cultivo de la papa y su origen	8
2.2	Importancia del cultivo de la papa	10
2.2.1	Importancia a nivel mundial	10
2.2.2	Importancia a nivel de Latinoamérica.....	13
2.2.3.	Importancia a nivel español	16
2.3	Parámetros de calidad de la papa	17
2.3.1.	Aptitudes culinarias de la papa	19
2.4	Imágenes Hiperespectrales (HSI).....	20
2.4.1.	Base Teórica.....	23
2.4.2.	Ventajas y desventajas	25
2.4.3.	Métodos de adquisición (escaneo) y detección de imágenes hiperespectrales	25
2.4.4.	Instrumentación.....	27
2.4.5.	Etapas del proceso de análisis.....	28
2.5	Quimiometría.....	31
2.5.1.	Análisis cualitativo.....	31
2.5.2.	Análisis cuantitativo.....	32
2.6	Aplicaciones HSI en papa	34
3.	MATERIAL Y MÉTODOS.....	35
3.1	Diseño experimental.....	35
3.2	Material vegetal.....	35
3.3	Adquisición de imágenes hiperespectrales.....	36
3.4	Software.....	38
3.5	Procesado de imágenes.....	38
3.6	Análisis quimiométrico: análisis multivariante.....	40

4. RESULTADOS Y DISCUSIÓN	43
4.1 Modelos de calibración para la clasificación de la papa según su aptitud industrial	43
4.2 Validación externa del modelo de calibración	46
5. CONCLUSIONES	49
6. BIBLIOGRAFÍA	50

ÍNDICE DE TABLAS

Tabla 1. Aporte nutricional de la papa en comparación con las pastas y el arroz. Fuente: (Muñoz, 2014).	9
Tabla 2. Superficie de tierra agrícola (%), superficie de papa en LA respecto al total agrícola (%), área (ha), rendimiento (t/ha) y producción anual de papa (TM) en LA. Fuente: (FAO, 2019; Gardi et al., 2014; Garzón, 2007).	14
Tabla 3. Principales diferencias entre la técnica de imagen, espectroscopía e imagen hiperespectral. Fuente: adaptada de (Sun, 2010).	23
Tabla 4. Ventajas y desventajas de la aplicación de la técnica de imágenes hiperespectrales a la industria alimentaria. Fuente: (ElMasry et al., 2007; Sun, 2010; Wu & Sun, 2013).	25
Tabla 5. Métodos de adquisición (escaneo) y detección de imágenes hiperespectrales. Fuente: adaptado de (Castro, 2015; ElMasry et al., 2012; Sun, 2010; Wu & Sun, 2013; Xiong et al., 2014).	26
Tabla 6. Variedades de papa empleadas en el estudio.	36
Tabla 7. Pretratamientos utilizados en los modelos desarrollados.	43
Tabla 8. Características de cada uno de los modelos.	44
Tabla 9. Resultados de la sensibilidad y especificidad de los 3 mejores modelos y los datos sin pretratamiento.	46
Tabla 10. Resultados de la validación externa del modelo de calibración.	47

ÍNDICE DE FIGURAS

Figura 1. Aporte (%) de una porción de 175 g de papa (2 tubérculos medianos) al total diario requerido. Fuente: (Muñoz, 2014).	9
Figura 2. Tendencia de producción de papa en el Mundo (2007-2017). Fuente: (FAO, 2019).	10
Figura 3. Porcentaje de producción de papa en los diferentes continentes. Fuente: (FAO, 2019). .	11
Figura 4. Principales países productores de papa en el año 2017. Fuente: (FAO, 2019).	11
Figura 5. Tendencia de producción de papa en los países menos desarrollados. Fuente: (FAO, 2019).	13
Figura 6. Distribución económica (%) y de volumen (%) de papa en España. Fuente: (Ministerio de Agricultura y Pesca, 2018).	17
Figura 7. Ilustración esquemática del cubo de imágenes hiperespectrales. Fuente: Adaptada de (Xiong et al., 2014).	20
Figura 8. Comparación entre imágenes hiperespectrales (derecha) y multispectrales (izquierda). Fuente: adaptado de (Durán, 2011).	21
Figura 9. Espectro electromagnético. Fuente: (López-Maestresalas, 2016).	22
Figura 10. Representación de un hipercubo. Fuente: (Ayala, 2018).	24
Figura 11. Esquema de un sistema de adquisición de imágenes hiperespectrales. Fuente: (W. Castro, 2015).	28
Figura 12. Pasos a seguir en el análisis de imágenes hiperespectrales. Fuente: elaboración propia.	29
Figura 13. Representación gráfica del despliegue (unfolding) de un hipercubo. Fuente: (Dorrepaal et al., 2016).	31
Figura 14. Equipo de adquisición de imágenes hiperespectrales: equipo real (a); esquema representativo (b) Fuente: adaptado de (Castro, 2017; López-Maestresalas et al., 2016)	37
Figura 15. Esquema del funcionamiento del espectrógrafo en el sistema de adquisición hiperespectral. Fuente: adaptado de (Durán, 2011).	38
Figura 16. Selección de píxeles (a) y sus respectivos espectros (b). Fuente: elaboración propia. ..	40
Figura 17. Efecto de los pretratamientos en los espectros de reflectancia: a) Sin pretratamiento, b) Smoothing, c) SM + SNV + MC y d) SM + 1ª D + MC.....	45
Figura 18. Representación gráfica de los resultados de predicción que ejecuta el modelo sobre la matriz de calibración externa.	47
Figura 19. Imagen química de las muestras utilizadas en la validación externa.	48

1. INTRODUCCIÓN Y OBJETIVO

De acuerdo a la cantidad de papa (*Solanum tuberosum* L.) producida a nivel mundial, este tubérculo es el quinto alimento básico del mundo, seguido del arroz, el trigo, el maíz y la caña. En el año 2017 la producción fue de 388.190.674 toneladas a nivel mundial (FAO, 2019). Se produce en la mayoría de los continentes, exceptuando la Antártida. Esto lo convierte en uno de los cultivos alimentarios más importantes del mundo, presentando un drástico crecimiento de producción en las últimas décadas, principalmente en los países en vías de desarrollo y se prevé que en el 2020 la producción de la papa adquiera un incremento especial, superando en área de producción a todos los demás cultivos alimentarios. El incremento de la producción se debe, al aumento de la población y el aumento del consumo per cápita, que corresponde a más de 70 kilogramos por persona al año. (Birch et al., 2012; FAO, 2006).

Dadas las estadísticas de crecimiento en la demanda del mercado, la cantidad de producto que se procesa es cada vez mayor y su clasificación se ha vuelto un punto estratégico para cumplir con los objetivos de producción en las industrias agroalimentarias, por ello muchas plantas de procesamiento se encuentran en la búsqueda continua de procedimientos eficientes y económicamente rentables que les permitan desarrollar las líneas de producción (Ji et al., 2019). La calidad de consumo de la papa es determinada por el valor que le da el consumidor a una serie de cualidades. En este aspecto, generalmente se realiza una distinción entre las cualidades del producto crudo y el producto cocido o frito. Entre las cualidades del producto cocido figuran entre otras, el color, la consistencia, la harinosidad, y el sabor. En los fritos las cualidades más importante son el color y la textura (Valdunciel, 2008).

Los requerimientos de calidad de la papa se basan en su destino final (papas fritas, congeladas, fécula, etc.), teniendo en cuenta que no todas las variedades proporcionan la misma calidad, las industrias exigen a los agricultores variedades determinadas por el proceso de transformación industrial. Poco a poco el número de variedades en el mercado ha ido aumentando y el consumidor va apreciando no sólo la forma del tubérculo, textura y color de la piel o facilidad de pelado, sino también la diferencia de calidad culinaria. Por lo que se suele especificar en el envase el nombre de la variedad y el tipo de consumo a que mejor se adapta (frito o cocido). Por todo ello es muy importante para la industria conocer si la papa que adquiere es la óptima para el destino final (Borruey et al., 1999; Valdunciel, 2008).

Tradicionalmente la clasificación de las papas se realiza por medio de análisis en laboratorio, con ensayos en su mayoría destructivos que requieren largo tiempo y por medio de inspectores de calidad, sin embargo los problemas derivados de esta clasificación incluyen la subjetividad de los seres humanos, la inconsistencia y, los altos costes, por lo que los resultados se ven fácilmente influenciados por el medio en el que se desarrolla la actividad (Su et al., 2017). Desde un punto de vista logístico y económico resulta atractivo determinar la aptitud de transformación industrial de las papas de manera rápida y eficiente, para cumplir con los requerimientos de calidad que solicitan los consumidores. Los métodos no destructivos proponen una solución rápida y útil para las industrias agroalimentarias, ya que permiten la medición simultánea de datos físicos sin destruir la muestra, además permiten obtener información cualitativa y cuantitativa sin necesidad de realizar análisis separados (El-Mesery et al., 2019). La aplicación de imágenes hiperespectrales podría permitir el desarrollo de un método no destructivo rápido y confiable para determinar diferentes características de las papas (López-Maestresalas et al., 2016). Actualmente existen diversos estudios de análisis de imagen hiperespectrales para una gran cantidad de productos del mundo agrario. Muchos de estos estudios utilizan el infrarrojo cercano ya sea para identificar daños o enfermedades, o bien para un análisis cuantitativo de los componentes de las muestras (Gowen et al., 2007).

En base a lo anterior, el objetivo de este estudio es evaluar la funcionalidad de imágenes hiperespectrales en el rango infrarrojo cercano como herramienta de clasificación de distintas variedades de papa en función de su aptitud de fritura y cocción.

2. REVISIÓN BIBLIOGRÁFICA

2.1 Cultivo de la papa y su origen

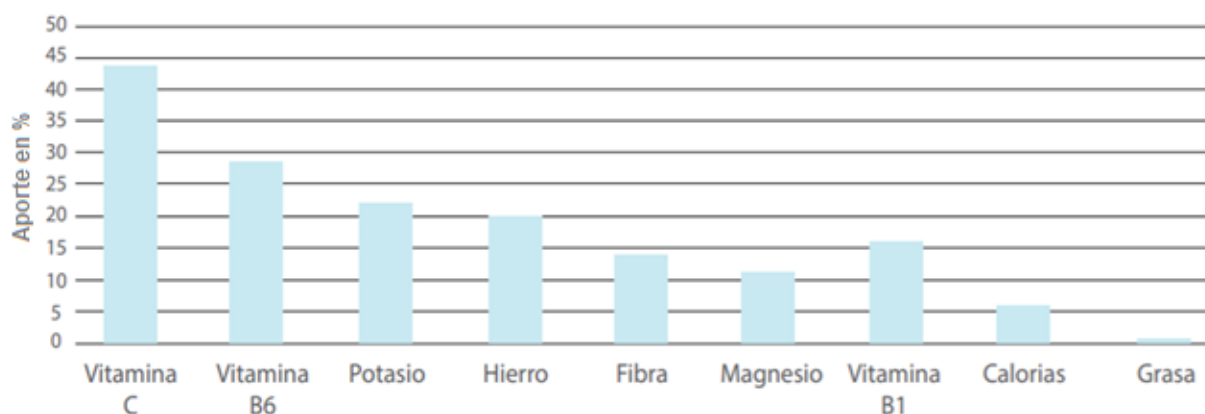
La papa es una planta perteneciente a la familia de las solanáceas, que por sus tubérculos comestibles es cultivada en todo el mundo. El cultivo requiere temperaturas de 15 a 20°C para la tuberización (formación de tubérculos) y crecimiento. Para un óptimo rendimiento de estas características a lo largo de su ciclo de producción es necesario que exista una variación, entre la temperatura diurna y la nocturna, de por lo menos 10°C. Sus capacidades de adaptación a diferentes climas, suelos y condiciones ambientales, convierten este cultivo en un alimento universal (Birch et al., 2012; FAO, 2006; INIA, 2017; Román & Hurtado, 2002).

En la formación del tubérculo inciden muchas condiciones climáticas, por ejemplo, la luminosidad que reciben las plantas durante el día incide en la formación de los diferentes tipos de azúcares que pasan a formar parte de los tubérculos. Al igual, la precipitación incide en el crecimiento de los tubérculos, donde el requerimiento es de 600 mm, distribuida en todo su ciclo vegetativo (Román & Hurtado, 2002). El ciclo vegetativo de la papa es aproximadamente de 80 días después de la brotación hasta la máxima acumulación de materia seca que corresponde a floración y continúa con la remoción de materia seca hacia los tubérculos, que aumentan su crecimiento hasta 130- 140 días después de la plantación. En la producción, el nitrógeno es uno de los nutrientes de mayor impacto, y se concentra principalmente en los tubérculos, estimándose que el 80% del nitrógeno adsorbido se vuelve a encontrar en ellos (INIA, 2017).

La papa es una buena fuente de energía, por su alto contenido de almidón, como se muestra en la Tabla 1, sus carbohidratos contienen menos calorías y grasas que otras fuentes de estos compuestos, como son el pan, las pastas o el arroz. Además, posee un alto contenido de vitaminas y minerales, detallado en la figura 1. Otros aportes nutricionales beneficiosos para la salud humana, son el contenido de fenoles, flavonoides y carotenoides (Muñoz, 2014).

Tabla 1. Aporte nutricional de la papa en comparación con las pastas y el arroz. Fuente: (Muñoz, 2014).

Característica	Papa (175 g)	Pasta (230 g)	Arroz (180 g)	Unidad
Energía (Calorías)	126	198	248	Kcal
Carbohidratos	27	43	56	g
Grasa	0,17	1,15	2,99	g
Proteínas	3,15	6,9	4,68	g
Fibra	2,1	2,07	0,18	g
Vitamina C	0,5	0	0	mg
Vitamina B6	0,58	0,023	0,13	mg

**Figura 1.** Aporte (%) de una porción de 175 g de papa (2 tubérculos medianos) al total diario requerido. Fuente: (Muñoz, 2014).

La papa se originó, hace 8000 años en las montañas de los Andes en América del Sur, en la frontera entre Bolivia y Perú, se llevó a Europa en el siglo XVI y de aquí siguió avanzando hacia otras partes del mundo, hoy se cultiva en más de 163 países, desde la planicie de Yunnan en China y las tierras bajas subtropicales de la India, hasta las montañas ecuatoriales de Java y las estepas de Ucrania, convirtiéndola en uno de los cultivos con mayor difusión a nivel mundial (Birch et al., 2012; FAO, 2006; USIL, 2019).

Durante cientos de años y como parte de su evolución, la papa se ha cruzado con otras plantas silvestres del mismo género, lo que ha dado como resultado a más de 5 500 variedades con diferencias en color, tamaño y sabor. Existen 235 especies de papa reconocidas actualmente, divididas en especies cultivadas y silvestres, por ello algunos botánicos no restringen la biodiversidad de la papa solo a las especies cultivadas ya que tiene más especies silvestres afines que cualquier otro cultivo (Luján, 1996; Román & Hurtado, 2002; USIL, 2019).

2.2 Importancia del cultivo de la papa

2.2.1 Importancia a nivel mundial

La papa, es consumida por más de mil millones de personas en el mundo. En términos de seguridad alimentaria, es un cultivo crítico, teniendo en cuenta el crecimiento poblacional y el aumento de la tasa de hambre. Según la FAO, a nivel mundial en el año 2017 se cosecharon 19 302.642 hectáreas, con un promedio de rendimiento de alrededor de 17 toneladas por hectárea (t/ha). La figura 2 detalla la tendencia general en la producción de la papa, del año 2007 a 2017 muestra un aumento del 23%, de 314 millones de toneladas (Mt) a 388 Mt. El aumento en la producción de la papa, se le atribuye tanto al aumento de los rendimientos como al de las áreas cultivadas (Birch et al., 2012; FAO, 2019).

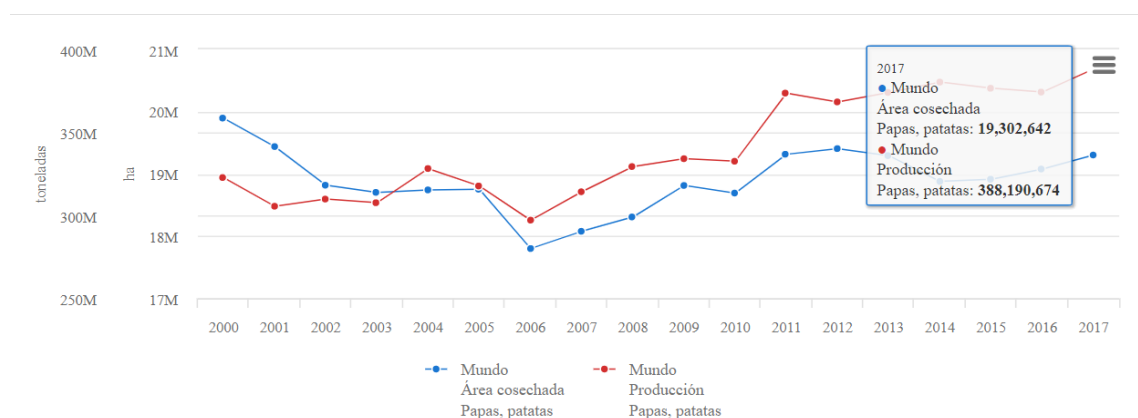


Figura 2. Tendencia de producción de papa en el Mundo (2007-2017). Fuente: (FAO, 2019).

Por su capacidad de adaptación y sus características nutricionales, la papa es producida en todo el mundo, siendo Asia y Europa, los continentes con mayor representación a nivel mundial (figura 3). Los cinco países con mayor producción son: China, India, Rusia, Ucrania y los Estados Unidos, estos países generan el 57% de la producción mundial de papa, y en las últimas décadas, China e India incrementaron la producción de tubérculos en 109% y 145%, respectivamente (figura 4) (FAO, 2019).

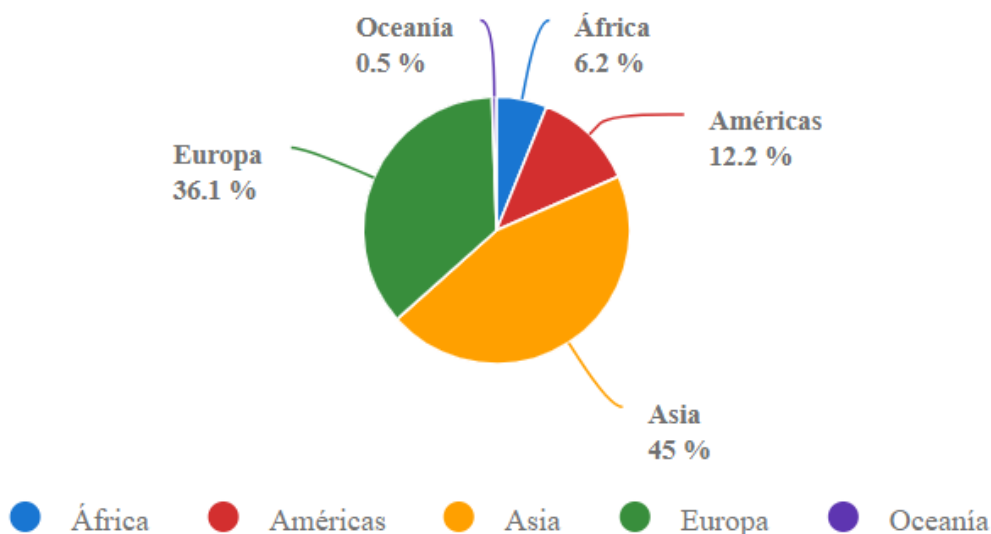


Figura 3. Porcentaje de producción de papa en los diferentes continentes. Fuente: (FAO, 2019).

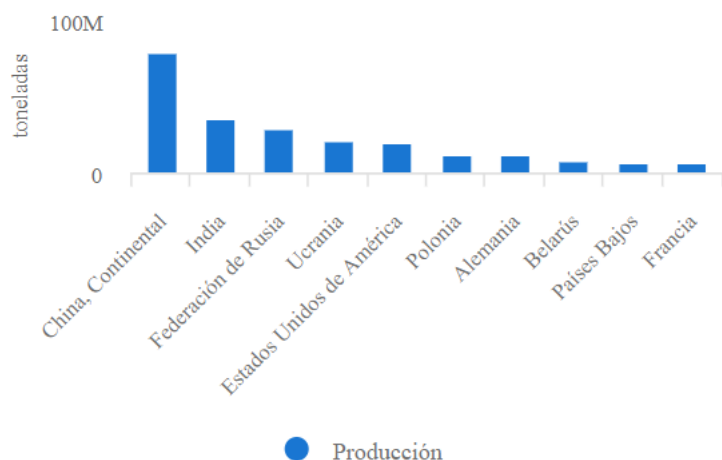


Figura 4. Principales países productores de papa en el año 2017. Fuente: (FAO, 2019).

En Europa, Australia y Estados Unidos de América, la mayoría de las papas se procesan fritas, enlatadas, deshidratadas y en puré. En el año 2013, los ingresos generados por la venta de papas fritas en los Estados Unidos de América fueron de 9.000 millones de dólares. Además, el valor de las exportaciones estadounidenses de papas fritas congeladas aumentaron más de 200 millones de dólares, del año 2006 al 2010 (H.Bohl et al., 2010). En el año 2017, China produjo un 395% más de papa, en comparación con los Estados Unidos de América, lo que posiciona a China en el país más productor de papa en el mundo, concentrando un poco más del 23% de la producción mundial, seguido por India y Rusia (FAO, 2019).

Países desarrollados como Canadá y Australia mostraron un considerable aumento de las importaciones de papas fritas a partir del año 2006, del 435% y 558%, respectivamente. Se notó una tendencia similar para algunos de los países en desarrollo como Brasil (229%), México (177%), y China (241%) (Faulkner, 2012).

Debido a la tendencia de crecimiento que ha mostrado la producción y el mercado de la papa, este cultivo se ha convertido en una actividad de importancia mundial. Por lo anterior el procesamiento de la papa está orientado a procesos altamente industrializados y tecnológicamente avanzados. La calidad de los productos y el éxito económico de esta industria dependen en gran medida de las materias primas de papa utilizadas y disponibles. Esto significa que la disponibilidad de cultivos adecuados y la calidad en sus diversos aspectos son de suma importancia para la industria de procesamiento de papa (Keijbets, 2008). China, al ser el mayor productor de papa en el mundo, desarrolló en el año 2008 las dos plantas más modernas para el procesamiento de papas fritas, bocadillos y almidón de papa. En el 2003 estas instalaciones procesaron alrededor de 1,26 millones de toneladas de papa. India, presenta otro de los mercados más emergentes y en el año 2007 inauguró una planta de procesamiento de papas fritas de última generación con una capacidad de 40.000 toneladas (Jansky et al., 2009; Keijbets, 2008).

En el año 2007 se registraron tres empresas norteamericanas y dos europeas que acaparan el 75% de la producción de la papa, actualmente estas cinco compañías operan en los cinco continentes, incluyendo a China y la India, que manejan gran parte del comercio internacional de productos procesados de papa, principalmente congelados y deshidratados. Más del 60% de las papas procesadas se utiliza para elaborar papas en bastones prefritas. La producción y el procesamiento continúan en crecimiento lo que ha generado un aumento de las tierras dedicadas a la producción y mejoras en el rendimiento de papa por hectárea (Keijbets, 2008). Adicionalmente las debilidades en las tecnologías de almacenamiento y transporte deben abordarse, ya que son las principales limitaciones para el desarrollo saludable de la industria de la papa en el mundo (Jansky et al., 2009).

En los últimos 20 años, los países en desarrollo han aumentado su participación en la producción, lo que indica que existe una tendencia de aumento en la producción de papa, en África y América Latina (figura 5). Según la FAO, estas regiones han estado más dispuestas a aceptar la introducción de nuevas variedades, mejorar la genética junto a una producción más eficiente. Existen estudios en varios países de África que muestran inversiones considerables en programas de distribución y almacenamiento para los tubérculos que ofrecen una base positiva para mejorar los niveles de productividad. La implementación de estas estrategias, combinadas con prácticas de producción cultural (semilla mejorada) dará como resultado grandes cifras de papas producidas y comercializadas (AGSI/FAO, 2001; FAO, 2017; USIL, 2019).

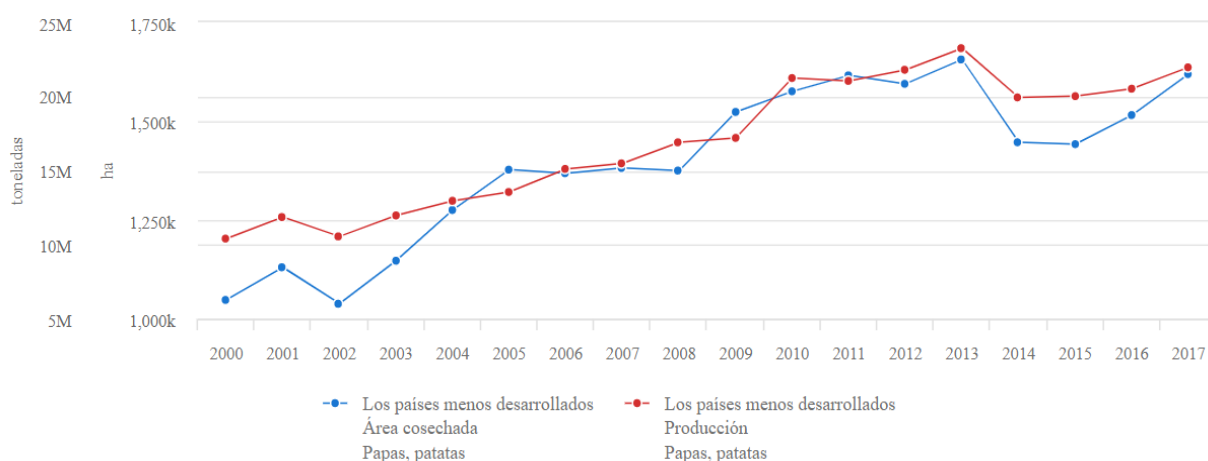


Figura 5. Tendencia de producción de papa en los países menos desarrollados. Fuente: (FAO, 2019).

2.2.2 Importancia a nivel de Latinoamérica

La papa en Latinoamérica (LA) es un cultivo que tiene mucha importancia tanto económica como social, ya que constituye la base alimentaria de varios países. Especialmente en las zonas altas de los Andes donde su aporte en la seguridad alimentaria es fundamental para los agricultores y las familias que la producen (Devaux et al., 2014).

La papa presenta una amplia distribución en Latinoamérica, desde el Sur hasta regiones del Norte. A lo largo del continente presenta especies cultivadas y silvestres, con una gran complejidad y diversidad en su composición genética. Los países que presentaron mayor producción en el año 2017 (Tabla 2) son Perú, Colombia, Brasil, y Argentina (FAO, 2019) y los que presentan mayor diversidad en especies y variedades son Perú y Bolivia (Garzón, 2007).

Los países que más superficie agrícola utilizan para producir papa son Perú, Bolivia, Chile y Colombia, entre un 2,0% y 6,2% de la superficie agrícola total. El resto de los países ocupan superficies menores a 2% de su suelo agrícola para producir papa, pero donde el volumen se mantiene elevado es en Brasil y Argentina, con producciones mayores a 2.000.000 de toneladas de producción anual, debido a sus elevados rendimientos mayores a las 22 t/ha (Tabla 2). En algunas regiones de Brasil, según la época de siembra (mayo) los rendimientos pueden superar las 35 t/ha (FAO, 2019; Gardi et al., 2014; Garzón, 2007).

Tabla 2. Superficie de tierra agrícola (%), superficie de papa en LA respecto al total agrícola (%), área (ha), rendimiento (t/ha) y producción anual de papa (TM) en LA. Fuente: (FAO, 2019; Gardi et al., 2014; Garzón, 2007).

País	Tierra Agrícola (%)	Área papa/Total agrícola (%)	Área Cosechada (ha)	Rendimiento (t/ha)	Producción (TM)
Argentina	12	0,2	77753	32	2491705
Bolivia	4	4,4	178144	5,8	1045291
Brasil	8	0,2	118130	30,9	3655069
Chile	2	3,3	54082	26,3	1426479
Colombia	5	2,0	162879	22,7	3706563
Cuba	38	0,1	6765	21,7	147044
Ecuador	10	1,8	29532	12,7	377243
Paraguay	10	0,01	260	14,4	3760
Perú	4	6,2	310400	15,3	4776294
Uruguay	25	0,1	4584	19,8	90772
Venezuela	4	1,0	11000	18,1	200000

TM: Toneladas Métricas, ha: Hectáreas, t: Toneladas.

Existe una gran variabilidad de los rendimientos de papa obtenidos en los diferentes países de LA (Tabla 2), esta variabilidad se atribuye a las características biofísicas, tipos de especies, características socioeconómicas y buenas prácticas agrícolas propias de los sistemas productivos de cada región. Los sistemas de producción agrícola de la papa en LA se clasifican en sistemas de alta o baja inversión. Los de alta inversión utilizan como recurso principal semilla de alta calidad, mecanización agrícola, fertilizantes y hacen un control intensivo de plagas y enfermedades del cultivo. En cambio, los sistemas de baja inversión utilizan la semilla propia, es decir aquella que ha sido seleccionada en los mismos campos durante muchos años, utilizan labranza mínima o basada en tracción animal, reducido uso de fertilizantes minerales y además hacen un escaso uso de plaguicidas (Condori et al., 2016).

La disponibilidad de inversión económica de cada país o región define en gran medida la prioridad del tipo de producción de papa, es decir, si el destino final del cultivo irá directamente a la comercialización en mercados formales o supermercados, o bien si será un cultivo para el consumo directo, familiar o para la venta en mercados locales. En países como el Perú la papa representa el 11% del producto interno bruto (PBI) agropecuario, están involucrados más de 600.000 agricultores de 19 departamentos del país. En los últimos años el consumo de papa ha aumentado notablemente, en 2001 se consumían un promedio de 54 kg/per cápita, actualmente se consumen 87 kg/per cápita. Esta actividad agronómica es catalogada como la que genera más ingresos para las comunidades alto Andinas. Debido a esto cuando se presentan inconvenientes en la producción y cosecha se originan problemas como la migración, lo que altera el orden social y económico de la región (Maza, 2011).

En Perú se encuentra la sede principal del Centro Internacional de la Papa (CIP), catalogado mundialmente como uno de los mayores centros dedicados a la investigación científica de la papa, que cuenta con diferentes alianzas mundiales que agrupan un total de 15 centros de investigación, cuenta con oficinas y programas en cerca de 40 países de Asia, África y América Latina, la sede principal fue fundada en el año 1971. El CIP tiene como objetivo estratégico disminuir la pobreza y alcanzar la seguridad alimentaria en países en desarrollo mediante la investigación de alternativas sostenibles basadas en la ciencia y preservación de la biodiversidad del planeta y sus recursos naturales. Además posee un banco genético con unos 5.000 tipos diferentes de papa silvestre y cultivada y tubérculos andinos provenientes de Bolivia, Ecuador y Perú. Igualmente, el CIP se encarga de la producción de semillas de papa mejoradas para resistir a enfermedades, heladas y sequías (CIP, 2018; Rutherford, 2014).

Durante la última década, en los países latinoamericanos ha habido un rápido desarrollo de nuevos mercados para la papa. El crecimiento de los supermercados a nivel regional es una tendencia clara de este crecimiento, intrínseco a este desarrollo se originan nuevas exigencias en cuanto a la calidad, tiempo de vida en anaquel y diferenciación de nuevas variedades. La demanda de nuevas variedades con calidad de procesamiento ha sido parcialmente atendida con la liberación de variedades como la Única y María Bonita. Sin embargo, las empresas procesadoras aún utilizan variedades antiguas como Capiro, Perricholi o Canchán. En forma paralela, la comunidad de desarrollo y nutrición junto al CIP, han desarrollado y popularizado la biofortificación de la papa, la cual consiste en un mejoramiento nutricional realizado por cruzamiento sexual para

lograr mayores contenidos de minerales que son déficit en la población latinoamericana, como el zinc (Zn) y el hierro (Fe) a fin de aumentar la ingesta en las dietas en zonas donde existe desnutrición (CIP, 2012; Shimizu, 2014).

2.2.3. Importancia a nivel español

En el año 2017 España produjo alrededor de 2.239.470 toneladas de papa, esta cifra se encuentra estabilizada desde el año 2005 en torno a las 2.000.000 toneladas, con unas 70.000 hectáreas de cultivo, y ocupa el octavo puesto entre los países de la UE. El rendimiento medio del cultivo está en unas 31 t/ha, lejos de países como Holanda, Bélgica, Alemania, Francia o Gran Bretaña que superan las 40 t/ha de producción media (FAO, 2019).

La cantidad producida del cultivo de la papa ha experimentado un descenso, superando las 3.000.000 toneladas a comienzo del año 2000 y bajando a las 2.000.000 toneladas para el año 2017, lo que representa un descenso del 33.33%. Del mismo modo, el área cosecha del cultivo de la patata ha experimentado un notorio descenso, pasando de las 118 mil hectáreas en el año 2000 a las 70 mil hectáreas en el año 2017. Esta reducción se ha visto compensada con el aumento del rendimiento de producción, que del año 2000 al 2017 ha aumentado un 20% pasando de 21 a 31 t/ha (FAO, 2019). En los países desarrollados la tendencia del consumo de papa presenta una disminución debido al cambio de los hábitos de consumo, la fama que tiene de ser un alimento que engorda y los nuevos estilos de vida caracterizados por falta de tiempo para cocinar (Arce, 2016; Señán, 2010).

La producción española está muy por debajo de la demanda del mercado. Se importan anualmente unas 700.000 toneladas, de las que más de 70% procede de Francia. Para el año 2017 Castilla León fue la comunidad autónoma más productora (887.584 toneladas y 20.433 hectáreas cultivadas), seguida por Galicia (383.170 toneladas y 17.96 hectáreas cultivadas) y Andalucía (317.976 toneladas y 10.279 hectáreas cultivadas) (MAGRAMA, 2018). Galicia, es la comunidad con el consumo per cápita más alto registrado en el 2018, con una cantidad de 34,25 kilos per cápita. Esta cantidad supera la media nacional. Al contrario, la comunidad de La Rioja, posee un consumo per cápita inferior en un 32% a la media nacional (Ministerio de Agricultura y Pesca, 2018).

En los hogares españoles se dedica el 1,62% del presupuesto para la alimentación y bebidas a la compra de papas, alcanzando un gasto per cápita de 24,20€ y un consumo per cápita de 28,18 kg por persona al año. Del total de consumo per cápita, las papas frescas son las más consumidas, con una ingesta media de 20,48 kilos, luego las papas procesadas con una ingesta media de 1,34 kilos y, por último, las papas congeladas con una ingesta de 0,98 kilos. Por lo anterior en términos de volumen de producción las papas frescas representan el 72,7% y el 65,7% de la facturación total, posicionándose como el tipo de mayor porcentaje. En un orden descendiente le siguen las papas procesadas que representan un 23,9% del volumen y un 29,3% de la facturación. Por último, se encuentran las papas congeladas con un 3,5% en volumen y 5,0% en valor económico, tal como se detalla en la figura 6 (Ministerio de Agricultura y Pesca, 2018).

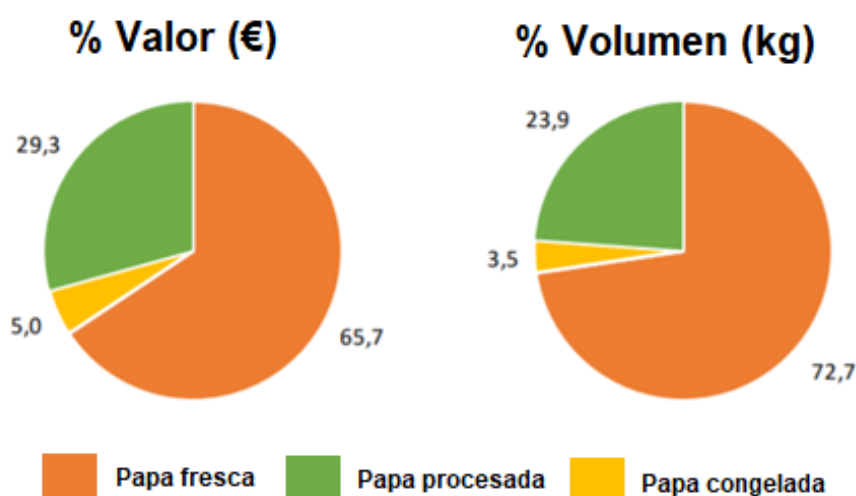


Figura 6. Distribución económica (%) y de volumen (%) de papa en España. Fuente: (Ministerio de Agricultura y Pesca, 2018).

2.3 Parámetros de calidad de la papa

En las industrias alimentarias debido a la necesidad de estandarizar la calidad de los productos industriales, se establecen especificaciones de la materia prima a utilizar. En el campo de la producción de papa, los rasgos de calidad deseados dependen del uso previsto. En cuanto a las características de calidad de la papa se debe hacer una distinción entre calidad externa del tubérculo y calidad interna (Andrade, 1997). Los parámetros de la calidad externa incluyen el tamaño, la forma, la textura y el color del tubérculo, así como la resistencia al estrés mecánico durante y después de la cosecha. La calidad interna de las papas está definida por parámetros

físico-químicos como el color de la pulpa, el contenido de materia seca y de almidón, el porcentaje de azúcares reductores, la susceptibilidad al pardeamiento enzimático y a las manchas negras y la decoloración después del cocido. El contenido de sólidos solubles en la papa es una de las características más importantes para el procesamiento industrial, ya que en la mayoría de procesos, contenidos altos son sinónimo de alto rendimiento; para los procesos industriales que involucren deshidratación como papa prefrita o papa tostada, se requiere un valor $>20\%$ (Esquivel et al., 2009). Muchas de estas características están estrechamente relacionadas y controladas por la genética. Además, se ha demostrado que todos estos parámetros también están relacionados con el estado nutricional tanto de la planta como de los mismos tubérculos. Otro factor que influye en las características de calidad es el manejo posterior a la cosecha (postcosecha) (Koch et al., 2020)

El almidón, además de ser una importante fuente de energía, tiene gran influencia en factores de calidad (Juszczak et al., 2012; Prada, 2008). De igual forma, existen criterios de calidad importantes en términos de seguridad alimentaria, como por ejemplo el potencial para formar compuestos cancerígenos como la acrilamida a partir de precursores durante la cocción y la acumulación de sustancias tóxicas como los glicoalcaloides (Koch et al., 2020).

Así como las industrias procesadoras establecen especificaciones de materia prima, los consumidores por su parte establecen parámetros de compra, como por ejemplo la apariencia externa de la papa como el color, forma, tamaño, brillo, estado sanitario, etc., además del sabor y el aroma. Estos parámetros establecidos por los clientes, los tienen muy en cuenta las industrias, al momento de establecer las especificaciones de materia prima. En general cada vez más, los parámetros de calidad de la papa están adquiriendo una mayor importancia en la industria y los programas de mejoramiento, ya que los consumidores demandan una mayor variedad y los minoristas desean comercializar productos que tengan ventajas comerciales distintivas. Sin embargo, como ocurre con muchos alimentos de campo, el sabor y la textura son difíciles de evaluar. Las evaluaciones son altamente subjetivas y requieren paneles sensoriales entrenados, como resultado, el sabor y la textura generalmente solo se evalúan en las etapas posteriores (Taylor, McDougall, & Stewart, 2007).

2.3.1. Aptitudes culinarias de la papa

El tratamiento culinario que se vaya a dar a la papa depende de la variedad y a su vez de la composición o características iniciales de la misma. Existen variedades óptimas para freír o para cocer, de tal forma que el tratamiento culinario influye directamente sobre las características de percepción sensorial y propiedades nutricionales. Los atributos de composición tienen que ver con la presencia o ausencia de ciertos componentes en el alimento como la humedad, el contenido de almidón y azúcares (Granados et al., 2014). Las principales características que determinan la mayoría de las propiedades organolépticas y la calidad del producto terminado son: la composición química, el contenido en azúcares reductores, el contenido en materia seca, la microestructura del tubérculo y el tamaño de la pared celular (Birch et al., 2012; Rubén et al., 2017).

Actualmente, la patata presenta diversas variedades obtenidas por selección y cruces. Esta diversidad proporciona un abanico de posibilidades del tubérculo permitiendo obtener diferentes comportamientos del producto en función del tipo de procesado al que se vaya a someter, ya que no todas las variedades tienen el mismo sabor, dan igual calidad de fritura, o mantienen su consistencia al cocer (Borruey et al., 1999; Gallego et al., 2006). La textura de los tejidos está positivamente relacionada con el contenido de materia seca, almidón, tamaño de los gránulos de almidón, gravedad específica y sólidos insolubles en alcohol que posean los tubérculos (Loyola et al., 2010).

Hoy en día se han establecido determinadas características, que unas veces son propias de la variedad y otras se consiguen con unas adecuadas prácticas culturales. Pero dado lo diverso de las utilidades industriales o culinarias es muy difícil que una variedad apta para un tipo de uso, lo sea también para otros distintos. El contenido en azúcares reductores óptimo para la industria de transformación en bastones para fritura debe ser menor de 2,5 g/kg en materia fresca ya que cuando este contenido supera el límite mencionado se obtienen productos de color oscuro y sabor amargo que son rechazados por el consumidor (Borruey et al., 1999). El contenido de materia seca se requiere entre 200-240 g/kg (20 – 25%) niveles superiores dan lugar a una textura dura y con contenidos inferiores a 200 g/kg, se obtienen productos blandos y con una mayor retención de aceite en el proceso de fritura (Gallego et al., 2006; Loyola et al., 2010). La papa óptima para cocción debe tener un contenido en materia seca entre el 17 y 20%, niveles superiores tienden a

ser más susceptibles a los moretones y los tubérculos se desintegran con más facilidad al cocer. El contenido de azúcares reductores debe ser menor al 0,2%, para evitar oscurecimiento (Valdunciel, 2008).

2.4 Imágenes Hiperespectrales (HSI)

La tecnología de imágenes hiperespectrales o HSI por sus siglas en inglés (Hyper Spectral Imaging) combina las ventajas de la visión por computador tradicional y la espectroscopía (Tabla 3); esta tecnología permite medir simultáneamente la variación espacial y espectral (figura 7) de una muestra (ElMasry, Sun, & Allen, 2013; Ji et al., 2019). Esta técnica se complementa con la espectroscopia ya que la visión por computadora por si sola es ineficaz en la identificación de objetos con colores similares, en las clasificaciones complejas, incapaz de predecir atributos de calidad (por ejemplo, composición química), y es ineficaz en la detección de defectos invisibles. Por el contrario, mediante las técnicas de espectroscopia, es posible obtener información acerca de los componentes de la muestra, pero no es fácil conocer la posición y/o ubicación de la información. De la combinación de las ventajas de estas técnicas surgen las imágenes hiperespectrales (Sun, 2010).

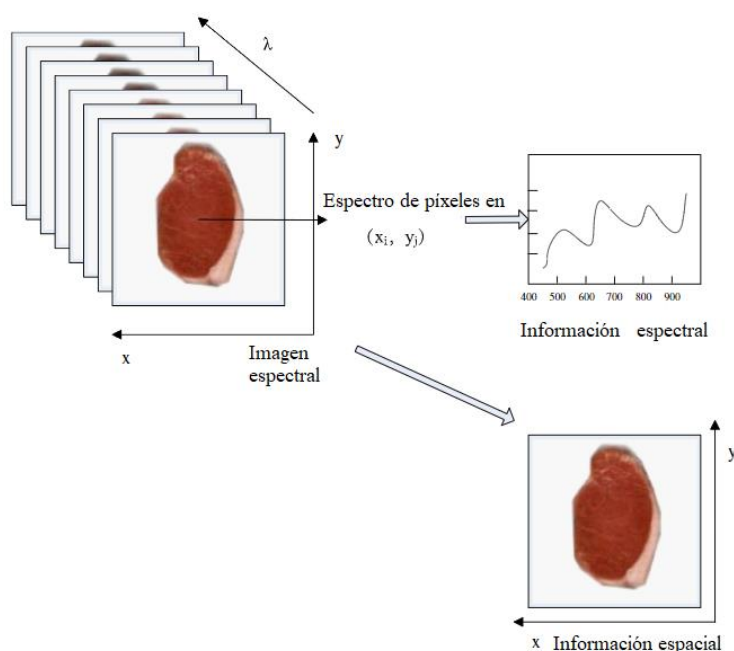


Figura 7. Ilustración esquemática del cubo de imágenes hiperespectrales. Fuente: Adaptada de (Xiong et al., 2014).

Las imágenes hiperespectrales están formadas por un gran número de bandas y estas siempre son contiguas. A diferencia de las imágenes multispectrales que están formadas por un menor (relativamente pocas) número de bandas y no son necesariamente contiguas unas con otras. Además, con una imagen multispectral se pueden obtener valores de intensidad en las longitudes de onda discretas en las que el sistema capte radiación, mientras que con una imagen hiperespectral lo que se obtiene es el espectro continuo o firma espectral del material analizado (Bernabé, 2016). En la figura 8, podemos observar las diferencias mencionadas anteriormente.

Teniendo en cuenta que la calidad de un alimento está constituida por diferentes componentes y características, el análisis por medio de imágenes hiperespectrales resulta altamente eficiente, ya que abarca gran parte de los componentes que la determinan y por medio de propiedades ópticas permite medir varias características simultáneamente sin destruir la muestra. Las propiedades ópticas se basan en la reflectancia, la transmitancia y la absorbancia en las regiones ultravioleta (UV), visible (VIS) e infrarrojo cercano (NIR) del espectro electromagnético medido con instrumentos espectrales (Sun, 2010).

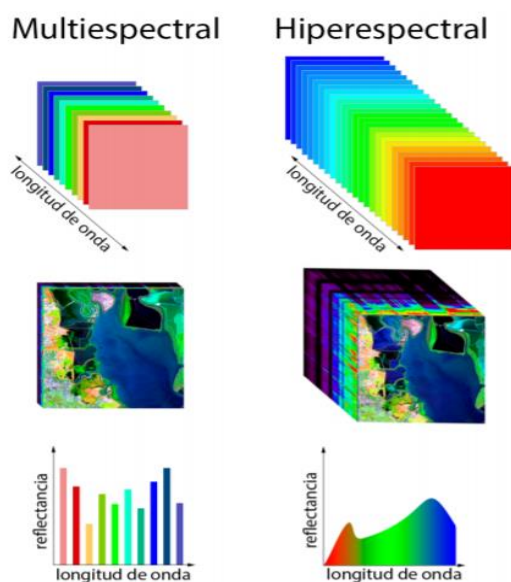


Figura 8. Comparación entre imágenes hiperespectrales (derecha) y multispectrales (izquierda). Fuente: adaptado de (Durán, 2011).

Del espectro electromagnético la región no detectable por el ojo humano más utilizada, para realizar estudios en la industria agroalimentaria es el rango infrarrojo (IR, infrared), que corresponde a las longitudes de onda entre 780 a 10^6 nm. A su vez esta región se divide en: infrarrojo cercano de 780 a 2500 nm (NIR), infrarrojo medio de 2500 a 50000 nm (MIR) e infrarrojo lejano de 50000 a 1000000 nm (FIR), tal como se muestra en la figura 9 (Lammertyn et al., 2000; López-Maestresalas, 2016). El NIR, es la región utilizada en este estudio.

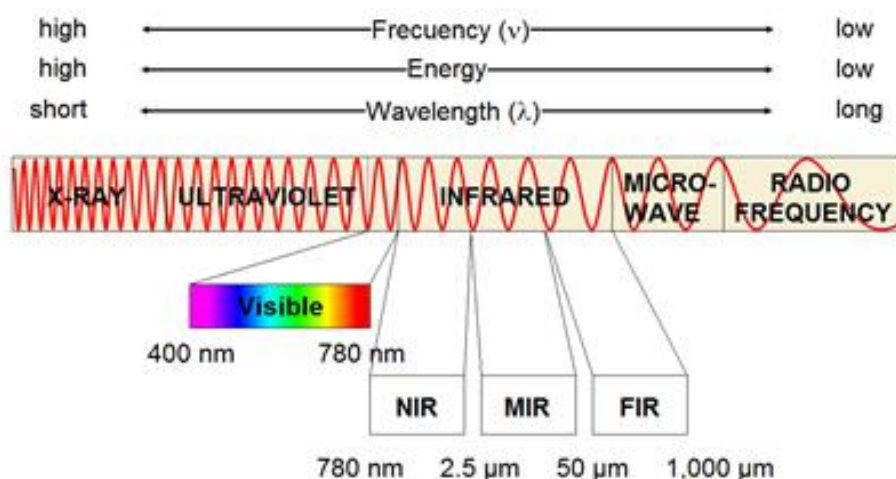


Figura 9. Espectro electromagnético. Fuente: (López-Maestresalas, 2016).

Las moléculas de la materia logran un estado vibracional de sus átomos, solamente cuando al absorber IR la frecuencia de la radiación electromagnética coincide con la frecuencia de vibración de los armónicos y las combinaciones de modos vibratorios fundamentales, es decir, para que ocurra el acoplamiento entre la radiación y la materia, la radiación debe de tener la energía precisa para satisfacer los requerimientos energéticos del material (Pasquini, 2018). La interacción de la radiación electromagnética con la materia causa un cambio en el momento dipolar de la molécula y cuando esto ocurre, la molécula pasa del estado de vibración fundamental a un estado de excitación de los modos vibratorios, aumentando la amplitud del movimiento vibratorio de los átomos del enlace. Los enlaces que vibran con gran energía y amplitud son los enlaces que contienen átomos de hidrógeno, dando lugar a bandas de gran intensidad. Por el contrario, en especies moleculares homonucleares como H_2 , O_2 y N_2 , no hay diferencias entre los estados vibracionales, por lo que no absorben radiación en el infrarrojo (Ozaki et al., 2018)

La variación espectral de la interacción de la radiación de la materia, para cientos de longitudes de ondas y diferentes puntos de muestreo, puede ser cuantificada por reflectancia (lo más común),

la cual es definida como el porcentaje de luz emitida por una muestra con relación a la luz total que incide en ella (Ji et al., 2019).

Tabla 3. Principales diferencias entre la técnica de imagen, espectroscopía e imagen hiperespectral. Fuente: adaptada de (Sun, 2010).

Características	Imagen	Espectroscopía	Imagen hiperespectral
Información espacial	Si	No	Si
Información espectral	No	Si	Si
Información multi - constituyente	No	Si	Si
Composición química	No	Si	Si

2.4.1. Base Teórica

El término imagen se refiere a una matriz de píxeles dimensionales de dos componentes espaciales $I \times J$, donde I son las columnas y J las filas. En el caso de imágenes en blanco y negro cada píxel tiene un valor de intensidad de grises. Mientras que en las imágenes a color son necesarias 3 imágenes para proporcionar los valores RVA (rojo, verde y azul) o RGB por sus siglas en inglés (red, green y blue) que dan la percepción de color al ojo humano. En las imágenes multi e hiperespectrales. Cada imagen tiene I columnas, J filas, además de las capas (H), dimensión que corresponde a la intensidad reflejada por cada píxel en las diferentes longitudes de onda medidas. Entonces una imagen digital puede describirse como una matriz cuyos índices de fila y columna identifican un punto en la imagen y cuyo valor coincide con el nivel de intensidad de la luz de ese punto (Grahn & Geladi, 2007; Rosenfeld & Kak, 1982).

Una imagen hiperespectral es una imagen que tiene varias capas superpuestas de imágenes monocromáticas, una por cada longitud de onda (λ) que se haya muestreado. Por ello, la matriz tiene un perfil tridimensional de píxeles (figura 10) con información espectral concreta en cada uno de ellos, lo que permite sacar un espectro completo de cada píxel (Mollazade, et al., 2012). El principio básico de la técnica de imágenes hiperespectrales es que todo material debido a su composición química y a su estructura física reacciona de una forma específica frente a la

radiación, ya sea reflejando parte de ésta en las diferentes longitudes de onda, absorbiendo, dispersando o emitiendo (Ayala, 2018).

En una imagen hiperespectral, la información se organiza en una matriz tridimensional (denominada hipercubo), en cuyas dos primeras dimensiones o ejes se almacena, mediante los conceptos de línea (X) y columna (Y), la información espacial y en la tercera dimensión (λ) la información espectral (Chang & Du, 2004; Cho et al., 2013; Xiong et al., 2014). Por tanto, un píxel situado en la coordenada (x, y) es un vector de bandas en el que cada elemento del vector corresponde a una de las diferentes longitudes de onda empleadas en el proceso de adquisición de la imagen y puede representarse en coordenadas longitud de onda (λ) vs reflectancia o intensidad (W. Castro, 2015).

El cubo hiperespectral tridimensional, o hipercubo, cubo de datos, volumen de datos, cubo espectral o volumen espectral, proporciona información del material estudiado. Esta información puede incluir observaciones físicas y geométricas de tamaño, orientación, forma, color y textura, así como información química/molecular como contenido de agua, grasa, proteínas y otros componentes unidos por hidrógeno (Sun, 2010). La característica única de cada material/componente es también conocida como firma espectral, la cual describe la radiación reflejada con respecto a la longitud de onda de dicha radiación (Bernabé, 2016).

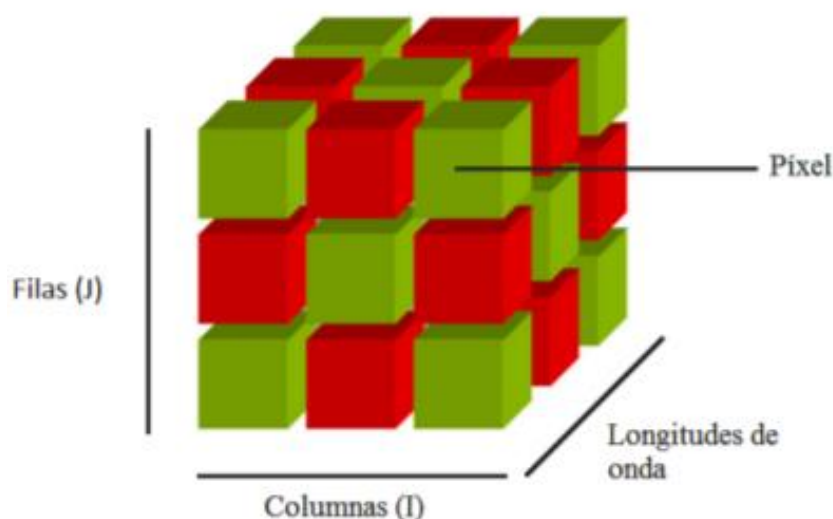


Figura 10. Representación de un hipercubo. Fuente: (Ayala, 2018).

2.4.2. Ventajas y desventajas

La principal ventaja de este tipo de imágenes es que cuando se trabaja con ellas se obtiene el espectro continuo o firma espectral del objeto de análisis. Por lo tanto, no se necesita un conocimiento previo de la muestra. En la tabla 4, se detallan las principales ventajas y desventajas de la técnica de imágenes hiperespectrales en la industria agroalimentaria.

Tabla 4. Ventajas y desventajas de la aplicación de la técnica de imágenes hiperespectrales a la industria alimentaria. Fuente: (ElMasry et al., 2007; Sun, 2010; Wu & Sun, 2013)

Ventajas	Desventajas
- Análisis libre de químicos.	- Requerimiento de gran velocidad del hardware para mejorar la adquisición y el análisis de datos.
- Ahorro de mano de obra, tiempo, reactivo, tratamientos de residuos.	- Necesidad de calibración precisa y modelos robustos.
- No invasión, ni destrucción de la muestra.	- Menores límites de detección en comparación a los métodos de análisis químicos.
- Delimitación múltiple de diferentes componentes dentro de una muestra en forma simultánea.	- No es aplicable a muestras líquidas u homogéneas, ya que se pierde su capacidad para relacionar la información espacial.
- Aplicaciones para análisis cualitativos y cuantitativos.	- Limitada penetración de la luz.
- Análisis relativamente rápido.	
- Visualización simultánea de la distribución espacial de diferentes parámetros de calidad.	
- Almacena abundante información para la evaluación de la calidad y seguridad alimentaria.	

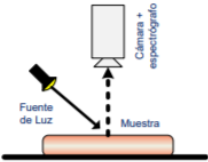
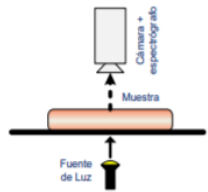

2.4.3. Métodos de adquisición (escaneo) y detección de imágenes hiperespectrales

Existen cuatro modos de adquisición de imágenes hiperespectrales en 3-D, conformados por diferentes elementos. Los métodos son: escaneo por puntos (point scanning), escaneo por líneas (line scanning imaging), escaneo de área o plano (area or plane scanning) y escaneo en una toma única (one single shot scanning). Los métodos de detección de las imágenes hiperespectrales son tres: reflectancia, transmitancia e interactancia (Castro, 2015; ElMasry et al., 2012; Sun, 2010;

Wu & Sun, 2013; Xiong et al., 2014). En la Tabla 5, se describe cada uno de los métodos de escaneo y detección, además se presenta un esquema de su funcionamiento.

Tabla 5. Métodos de adquisición (escaneo) y detección de imágenes hiperespectrales. Fuente: adaptado de (Castro, 2015; ElMasry et al., 2012; Sun, 2010; Wu & Sun, 2013; Xiong et al., 2014).

Método de adquisición	Descripción	Esquema
Escaneo por puntos (point scanning)	Se escanea un punto único en un píxel para proporcionar el espectro de este punto, y se escanean otros puntos moviendo el detector o la muestra a lo largo de dos dimensiones espaciales (X y Y). El formato de salida es BIP (Banda intercalada por píxel).	
Escaneo por líneas (line scanning imaging)	<p>En este método la cámara cuenta con una fila de detectores que registra una línea completa de una imagen con información espectral de cada píxel espacial en la línea. El cubo hiperespectral completo se obtiene desplazando la muestra. El formato de salida es BIL (Banda intercalada por línea).</p> <p>Debido a sus características de escaneo continuo en una dirección, el escaneo en línea es particularmente adecuado en sistemas de cintas transportadoras que se usan comúnmente en líneas de proceso de alimentos. Por lo tanto, el escaneo en línea es el método más popular para adquirir imágenes hiperespectrales para la inspección de calidad y seguridad de los alimentos.</p>	
Escaneo de área o plano (area or plane scanning)	<p>Se obtiene una imagen monocromática 2D (x, y), con información espacial completa a una sola longitud de onda a la vez. La exploración se repite en todo el rango de longitud de onda y da como resultado una pila de imágenes dentro de un área rectangular, el sistema codifica cada línea de la imagen en la primera banda seguida inmediatamente por la siguiente línea en la misma banda espectral, seguida por la segunda banda para todas las líneas, seguida por la tercera banda para todas las líneas, etc., permitiendo la generación de la imagen de toda la muestra.</p> <p>No es necesario mover ni la muestra ni el detector. Para este método es necesario generar una imagen por cada longitud de onda, por lo que es necesario un filtro para ajustar la longitud de onda deseada. Una desventaja del escaneo de área es que no es adecuado para realizar la inspección en tiempo real. El formato de salida es BSQ (Banda secuencial).</p>	
Escaneo en una toma única (one single shot scanning)	Este método registra información tanto espacial como espectral utilizando un detector de área grande con una sola exposición para capturar un espectro completo de la muestra, lo que lo hace muy atractivo cuando se requieren imágenes hiperespectrales rápidas. Sin embargo, todavía está en la etapa inicial de desarrollo y tiene resoluciones limitadas para la dimensión espacial y rangos estrechos para la dimensión espectral.	
Método de detección	Descripción	Esquema

Reflectancia	<p>El sensor captura la luz reflejada de la muestra desde el mismo lado en que es iluminada.</p> <p>Modo usado para la obtención de características de calidad externos (tamaño, forma, color, textura superficial y defectos externos).</p>	
Transmitancia	<p>El detector está ubicado en el lado opuesto de la fuente de luz, y captura la luz transmitida a través de la muestra que transporta información interna.</p> <p>Se utiliza generalmente para determinar la concentración de los componentes internos y defectos internos de objetos relativamente transparentes tales como pescado, frutas y verduras. Sin embargo, el modo de transmitancia tiene un nivel de señal bajo debido a la atenuación de la luz y se ve afectado por el grosor de la muestra.</p>	
Interactancia	<p>La fuente de luz y el sensor están situados en el mismo lado de la muestra y paralelos entre sí.</p> <p>Este modo puede detectar información más profunda de la muestra y tiene menos efectos de superficie en comparación con el modo de reflectancia. Además, el modo de interacción reduce la influencia del grosor, que es una ventaja práctica sobre la transmisión.</p>	

2.4.4. Instrumentación

La instrumentación de los sistemas de imágenes hiperespectrales es básica e importante para adquirir imágenes hiperespectrales confiables con alta calidad. La selección de los componentes de los instrumentos y el diseño de su configuración y calibración requieren una buena comprensión de la configuración y calibración del sistema (Wu & Sun, 2013). El sistema se compone de los siguientes elementos: una fuente de luz, un dispositivo de dispersión de longitud de onda, un detector de área y una plataforma de muestra (W. Castro, 2015; López-Maestresalas, 2016; Nicolai et al., 2007; Sun, 2010; Wu & Sun, 2013). En la figura 11, se muestra el esquema de un sistema de adquisición de imágenes hiperespectrales.

- **Fuente de luz:** generan luz para excitar o iluminar el objetivo. Las más utilizadas en los sistemas de imágenes hiperespectrales incluyen lámparas halógenas (fuente de iluminación de banda ancha usados en longitudes de onda visible (VIS) y el infrarrojo cercano (NIR); generalmente son de tungsteno), diodos emisores de luz (elementos emisores de luz, pequeños, de bajo coste, respuesta rápida, larga vida y bajo consumo de energía) y láseres (fuentes de luz monocromáticas principalmente usados para excitar los electrones en la muestra (W. Castro, 2015; Sun, 2010; Wu & Sun, 2013).
- **Dispositivo de dispersión de longitud de onda:** Tienen la función de dispersar la luz provista por el sistema de iluminación en diferentes longitudes de onda. Dentro de éstos

tenemos los espectrofotómetros de imágenes, filtros sintonizables acústico-ópticos AOTF (Acoustic-Optic Tunable Filters) y los filtros sintonizables de cristal-líquido LCTF (Liquid Crystal Tunable Filter) (W. Castro, 2015; Wu & Sun, 2013).

- **Detector de área:** la función de este dispositivo es cuantificar la información espacial y espectral en forma de luz (una vez que pasa por el dispersor de longitudes de onda) adquirida mediante la conversión de fotones incidentes en electrones a través de un circuito integrado. Este circuito puede ser un dispositivo de carga acoplada CCD (Coupled Charged Device) o un semiconductor complementario de óxido metálico CMOS (Complementary Metal Oxide Semiconductor). Ambos convierten la energía de radiación a señal eléctrica (Sun, 2010; Wu & Sun, 2013).
- **Plataforma de muestra:** en función del dispositivo de dispersión es posible que se requiera de una plataforma que desplace a la muestra misma a una velocidad constante mientras se realiza la captura de la escena por parte de la cámara (detector de área) (Gowen et al., 2007).

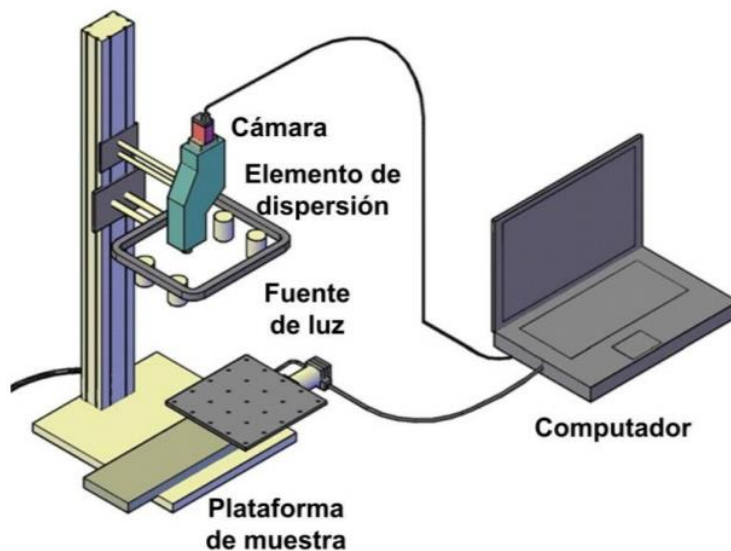


Figura 11. Esquema de un sistema de adquisición de imágenes hiperespectrales. Fuente: (W. Castro, 2015).

2.4.5. Etapas del proceso de análisis

Una vez elegido el método de detección (normalmente reflectancia), la configuración y la iluminación, se adquiere la imagen de una muestra que se almacena en varios archivos (png). En la figura 12, se ilustra el procedimiento de adquisición y análisis de imágenes hiperespectrales,

el cual se puede dividir en dos etapas: a) Adquisición y preprocesado, de acuerdo a los métodos de adquisición detallados en el apartado 2.4.3, de esta etapa resulta la información espectral; b) Modelización o quimiometría, en la cual se utilizan técnicas de análisis estadístico multivariante para correlacionar los parámetros analizados con la información espectral (W. Castro, 2015; Wu & Sun, 2013).

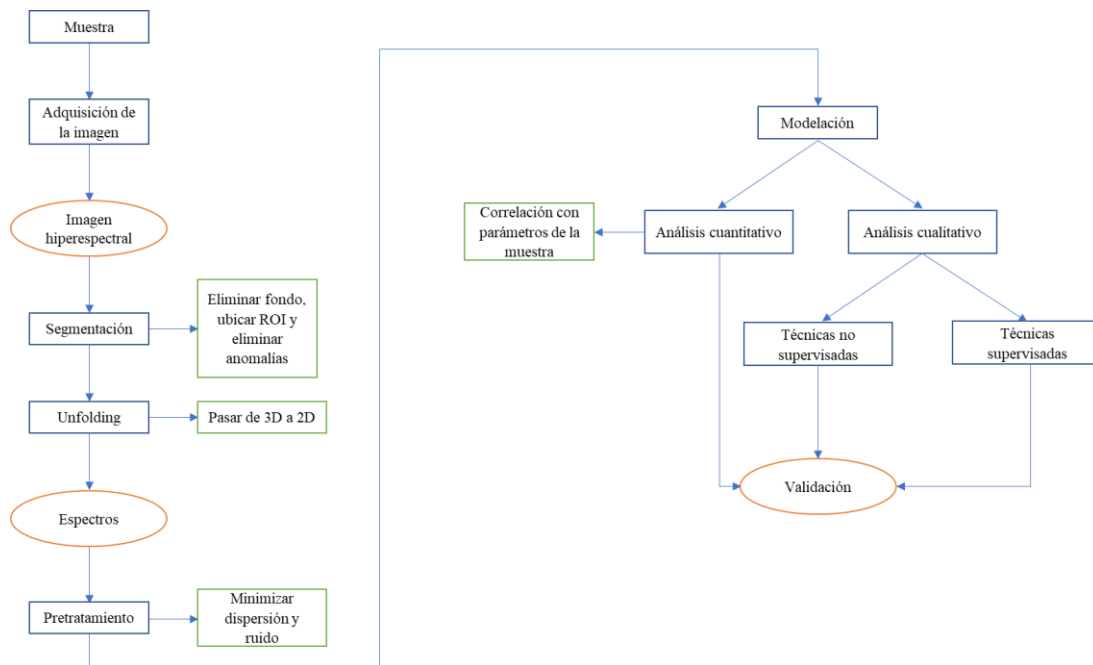


Figura 12. Pasos a seguir en el análisis de imágenes hiperespectrales. Fuente: elaboración propia.

Adquisición y preprocesado: permite obtener y comprimir la información para reducir el almacenamiento requerido, además de obtener una imagen hiperespectral corregida. En esta etapa se incluye la *segmentación* que se realiza con el objetivo de separar el fondo de la imagen analizada y ubicar las regiones de interés (ROI). El fondo genera espectros con mucho ruido y dificulta el resultado de los modelos multivariantes. Los métodos para eliminar el fondo son: histogramas basados en PCA, puntuaciones (scores) en los diagramas de dispersión de PCA, coeficientes específicos de λ , selección manual y K-means minimizando la suma de distancias entre cada observación y el centroide de su grupo o clúster (distancia euclídea). Otro paso de esta etapa es la *detección de anomalías* como píxeles muertos o líneas causadas por anomalías en el detector y picos irregulares. Cuando aparecen líneas enteras de píxeles muertos, se eliminan por completo y cuando son píxeles dispersos se interpolan con el valor de sus vecinos. Los Picos (*spikes*) son medidas erróneas de una longitud de onda, se deben detectar y eliminar para cada píxel. La detección de estos defectos y su tratamiento puede realizarse por diferentes vías, siendo la detección manual primordial para su tratamiento (Wu & Sun, 2013). Además, se realiza la

calibración de la reflectancia que tiene el propósito de normalizar la no linealidad del iluminante a lo largo del espectro y eliminar posibles ruidos constantes. Esto se realiza capturando un patrón que tenga un alto coeficiente de reflexión (máximo) en el rango que se está trabajando, llamado comúnmente “blanco” y otro patrón con el mínimo coeficiente de reflexión “negro” (Ivorra, 2015).

Cuando ya se obtiene la imagen hiperespectral corregida y se tienen las ROI se realiza el despliegue o **unfolding**, que consiste en pasar la información almacenada en el hipercubo (3D) a una matriz bidimensional (2D). Esto se hace para poder aplicar técnicas de análisis quimiométrico, ya que la mayoría de los algoritmos empleados están diseñados para trabajar en dos dimensiones. Los valores de intensidad espectral contenidos dentro del hipercubo no cambian ni el número de elementos del conjunto de datos. Los datos simplemente se reorganizan en dos dimensiones para ayudar en el análisis. Por lo anterior, la nueva matriz 2D (figura 13) provee la reflectancia de cada pixel en cada longitud de onda, donde el número de filas corresponde con el número de píxeles de la ROI de la muestra y el número de columnas con el número de longitudes de onda captadas (Dorrepaal et al., 2016).

A los espectros generados se les realiza un **pretratamiento** o **preprocesado** para minimizar la dispersión y el ruido. La dispersión se minimiza por medio de: Corrección de la variación normal estándar (Standard Normal Variate) –SNV, Corrección multiplicativa de la dispersión (Multiplicative Scatter Correction) –MSC y derivadas. El ruido espectral se minimiza por medio de: Suavizado (Smoothing), que consiste en aplicar una derivada de orden cero a los datos espectrales y Centrado medio (Mean centre), donde se calcula el espectro medio del set de datos y se le resta a cada espectro (Dorrepaal et al., 2016). Una vez realizado este proceso y todos los tratamientos previos, la información espectral de la muestra será apta para iniciar el proceso de análisis quimiométricos o análisis de datos.

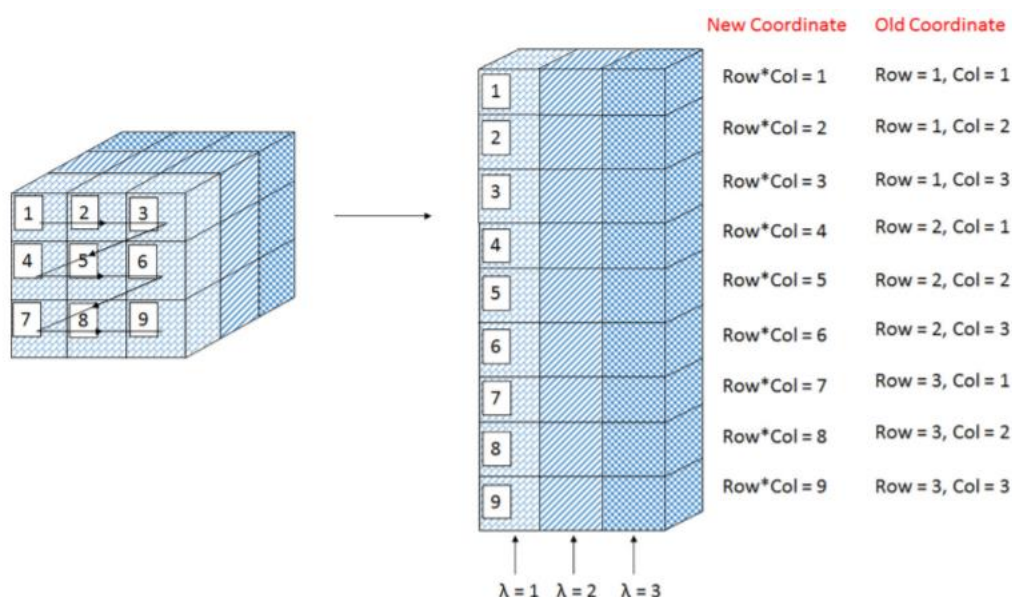


Figura 13. Representación gráfica del despliegue (unfolding) de un hipercono. Fuente: (Dorrepaal et al., 2016).

2.5 Quimiometría

Como se discutió anteriormente, una de las ventajas en el desarrollo de imágenes hiperespectrales es la gran cantidad de datos que comúnmente se extraen. Por ello se requiere el uso de la quimiometría, la cual se basa en un análisis multivariante para descomponer eficientemente una cantidad masiva de características en información útil y establecer una relación simple y más fácil de entender entre los datos de imágenes hiperespectrales y los atributos de interés de las muestras analizadas. El análisis multivariante se puede clasificar en análisis cualitativo y cuantitativo. En general, este análisis se implementa para analizar la información espectral y construcción de modelos de calibración/predicción para visualizar la información química oculta en las imágenes (ElMasry et al., 2012; Wu & Sun, 2013).

2.5.1. Análisis cualitativo

La clasificación cualitativa (también llamada reconocimiento de patrones), se basa en reconocer patrones o características intrínsecas a la muestra que posteriormente se asocian a una clase o categoría, agrupando la información con un cierto porcentaje de acierto. Incluye la clasificación no supervisada y la clasificación supervisada. El reconocimiento de patrones *supervisados*, también conocidos como análisis discriminantes, tiene como objetivo establecer un modelo para clasificar nuevas muestras desconocidas a clases conocidas previamente definidas (por ejemplo:

para cocción o fritura). Los algoritmos de clasificación multivariante supervisados típicos para el análisis de datos hiperespectrales incluyen análisis de discriminación lineal (LDA), análisis discriminante de mínimos cuadrados parciales (PLS-DA), redes neuronales artificiales (ANN), máquinas de vectores de soporte (SVM) y k-vecino más cercano (kNN). Como algoritmo de aprendizaje basado en instancias, kNN es quizás el más simple de todos los algoritmos de aprendizaje automático, donde un objeto se asigna a la clase por voto mayoritario de sus vecinos (ElMasry et al., 2012; Wu & Sun, 2013).

Por otro lado, las clasificaciones *sin supervisión* no requieren un conocimiento previo sobre las clases de los datos, por tanto, el número de clases es desconocido. Estos métodos buscan agrupar la información, para ello, intentan determinar fuentes de agrupamiento entre las muestras analizadas. Este tipo de análisis es a menudo muy útil para la evaluación preliminar de los contenidos de información en el conjunto de datos espectrales. Los algoritmos de clasificación multivariante no supervisados típicos para el análisis de datos hiperespectrales incluyen análisis de componentes principales (PCA), agrupamiento de K-medias cluster y agrupamiento jerárquico (ElMasry et al., 2012; Wu & Sun, 2013).

El análisis por PCA descompone los datos espectrales en varios componentes principales (PC) para caracterizar las direcciones más importantes de variabilidad en el espacio de datos de alta dimensión. La agrupación de K-medias cluster clasifica las muestras en K agrupaciones en las que cada muestra pertenece al grupo con la distancia mínima al centroide del grupo. La agrupación jerárquica está destinada a construir una jerarquía de agrupaciones que generalmente se presenta en un dendrograma (ElMasry et al., 2012; Wu & Sun, 2013).

2.5.2. Análisis cuantitativo

El modelo de regresión multivariante consiste en construir una relación entre los atributos físicos, químicos o biológicos deseados de una muestra analizada y su respuesta espectral, con propósitos explicativos o predictivos. La regresión multivariante puede ser lineal o no lineal. Los métodos de *regresión lineal* multivariante incluyen principalmente regresión lineal múltiple (MLR), regresión de componentes principales (PCR) y regresión de mínimos cuadrados parciales (PLSR) (ElMasry et al., 2012; Wu & Sun, 2013).

El análisis MLR establece una relación entre el espectro y los atributos deseados de la muestra analizada en forma de una ecuación lineal, es decir una relación entre una variable de respuesta y variables explicativas. MLR falla cuando el número de variables es mayor que el de las muestras y es fácil verse afectado por la colinealidad entre las variables. Los coeficientes de regresión de esta ecuación están determinados por el proceso de calcular el error mínimo entre los valores de referencia y los pronosticados.

El análisis PCR es un método multivariante de dos pasos: regresión por PCA y MLR. Primero, se realiza un PCA sobre datos espectrales, en lugar de las variables originales, las PC se usan como variables independientes en un MLR. Las ventajas del análisis PCR contra MLR es que el cálculo de PC hace que las variables independientes no estén correlacionadas y contengan menos ruido. La limitante del PCR es que los primeros componentes principales que se utilizan para el modelo de regresión no son necesariamente los más informativos con respecto a la variable de respuesta (ElMasry et al., 2012; Wu & Sun, 2013).

El análisis PLSR está emergiendo como la herramienta más robusta y confiable en el desarrollo de un modelo de calibración multivariante. Está construido en base al algoritmo lineal, por lo que obtiene un buen rendimiento cuando hay una relación lineal entre espectros y propiedades de objetos. Trata de predecir un conjunto de variables dependientes (concentraciones) a partir de un gran conjunto de variables independientes (información espectral). Esta predicción se logra extrayendo un conjunto de factores ortogonales llamados variables latentes (LV) que tienen el mejor poder predictivo (ElMasry et al., 2012; Wu & Sun, 2013).

Los métodos *no lineales* se utilizan cuando los datos espectrales y los atributos objetivo no tienen una relación lineal. Algunos de los métodos no lineales más comunes son: red neuronal artificial (ANN) y la regresión vectorial de soporte (SVR). El análisis ANN simula el comportamiento de las redes neuronales biológicas para el aprendizaje y el propósito de predicción. La red neuronal de alimentación multicapa es la técnica ANN más utilizada que tiene tres capas: de entrada, oculta y de salida para organizar las neuronas artificiales. El análisis SVM es un método basado en aprendizaje estadístico teórico que se basa en dos fases, la primera de entrenamiento, donde se les informa con múltiples ejemplos ya resueltos, en forma de pares (problema – solución). La segunda fase de uso para la resolución de problemas. En ella, las SVM se convierten en una “caja negra” que proporciona una respuesta (salida) a un problema dado (entrada).

2.6 Aplicaciones HSI en papa

En la industria alimentaria existen numerosas aplicaciones de imágenes hiperespectrales. En su mayoría son utilizadas para caracterizar cuantitativa y/o cualitativamente los componentes de una muestra, así como para detectar la presencia de defectos y enfermedades. A continuación, se nombran algunos estudios que contemplan la aplicación de esta tecnología en papa.

- Detección de costra común en papas usando un sistema de imagen hiperespectral en el rango infrarrojo (IR) (Dacal-Nieto et al., 2011a).
- Detección no destructiva de alteraciones internas (hollow heart) en papas usando imágenes hiperespectrales (Dacal-Nieto et al., 2011b).
- Detección de defectos externos en papas por tecnología de imágenes hiperespectrales y método de procesamiento de imágenes (Wenhao et al., 2014).
- Espectroscopía hiperespectral para la detección de la enfermedad del tizón temprano en plantas de papa en dos etapas de crecimiento diferentes (Atherton et al., 2015).
- Detección no destructiva de puntos negros (blackspot) en papas, con tecnología de imagen hiperespectral VIS-NIR y SWIR (López-Maestresalas et al., 2016).
- Espectroscopía de infrarrojo cercano e imágenes hiperespectrales para inspección de calidad no destructiva de papas (López-Maestresalas, 2016).
- Detección de papas magulladas utilizando la técnica de imagen hiperespectral basada en la transformación discreta ondícula (Wavelet transform) (Ji et al., 2019).

3. MATERIAL Y MÉTODOS

3.1 Diseño experimental

A continuación, se detalla el proceso seguido para evaluar la funcionalidad de imágenes hiperespectrales en el rango infrarrojo cercano para clasificar distintas variedades de papa en función de su aptitud culinaria de fritura y cocción.

Las imágenes hiperespectrales se obtuvieron de una rebanada de cada papa, se utilizaron 72 muestras de 18 variedades (4 muestras por variedad), 9 con aptitud de cocido y 9 con aptitud de fritura. Para la adquisición de las imágenes se utilizó un sistema hiperespectral de escaneo lineal sensible en el rango NIR (900-1700nm). Las imágenes fueron procesadas en MATLAB para obtener la información espectral y espacial de cada una de ellas. Se seleccionó de forma aleatoria el 75% de las muestras (54 imágenes, 1 imagen/tubérculo) para construir el modelo de clasificación y el 25% restante (18 muestras) se empleó para validar ese modelo. Luego, para realizar la calibración con la ayuda de HYPER-Tools v.2 se seleccionaron de forma manual 10 píxeles de cada imagen, obteniendo así 10 espectros por cada muestra. Con estos espectros se construyó la matriz de calibración (540 espectros) y se analizó mediante el método discriminante PLS-DA, con el fin de generar un modelo de clasificación de papas según su aptitud culinaria en base a sus diferencias espectrales.

3.2 Material vegetal

Para la ejecución de este estudio se utilizó una base de datos elaborada por el Grupo de Investigación Mecatrónica Agraria de la Universidad Pública de Navarra (UPNA). En la elaboración de la base de datos se utilizaron tubérculos de papa de 18 variedades diferentes (tabla 6): 9 variedades clasificadas con aptitud industrial para cocido y 9 variedades con aptitud industrial para frito; cosechadas en octubre de 2016. Se utilizaron 4 papas por variedad, sumando un total de 72 tubérculos. Todas las papas fueron suministradas por el Instituto Vasco de Investigación y Desarrollo Agrario (NEIKER-Tecnalia), y enviadas al Departamento de Ingeniería de la UPNA, donde se mantuvieron refrigeradas a 10°C hasta la realización de los análisis posteriores. En base a lo anterior se establecieron dos grupos de clasificación para el estudio:

- Papas con aptitud de transformación para fritura

- Papas con aptitud de transformación para cocción.

Tabla 6. Variedades de papa empleadas en el estudio.

Aptitud industrial Cocido		Aptitud industrial Frito	
Variedad	N	Variedad	N
Constance	4	Alicante	4
Jaerla	4	Bordura	4
Laudine	4	Babylon	4
Levantina	4	Corsica	4
Madeleine	4	Hermes	4
Marfona	4	Lady Amarilla	4
Montana	4	Lyoness	4
Rosa Gold	4	Honorata	4
Rudolph	4	Ottawa	4

Para elaborar la base de datos se caracterizaron las papas, para ello se enumeraron las de cada variedad y una vez identificadas, se pesaron y midieron individualmente. La adquisición de las imágenes se realizó empleando como material rebanadas de papa de 1 cm de espesor. Se utilizó una rebanada por tubérculo, sumando un total de 72 muestras. En la preparación de las muestras se empleó una cortadora de fiambre Braher (Modelo USA280) obteniendo una rebanada de cada tubérculo.

3.3 Adquisición de imágenes hiperespectrales

La adquisición de las imágenes hiperespectrales para desarrollar la base de datos, se llevó a cabo en el Laboratorio de Imagen de la Escuela Técnica Superior de Ingeniería Agronómica y Biociencias (ETSIAB), en la UPNA. Las rebanadas de papa se escanearon con un equipo de análisis de imagen (figura 14) compuesto por los siguientes elementos: un sistema hiperespectral (sensible en el rango NIR de 900 a 1700 nm), una plataforma móvil, una fuente de iluminación y un ordenador.

Las especificaciones técnicas del sistema hiperespectral utilizado en este estudio, son las siguientes: el equipo consta de una cámara Xeva 1.7-320-100Hz (Xenics, Leuven, Bélgica), con un sistema de escaneo lineal (figura 15) sensible en el rango NIR de 900 a 1700nm, con un detector InGaAs de resolución 320 x 256 píxeles y conexión USB, acoplada a un espectrógrafo

ImSpector N17E (Specim, Spectral Imaging Ltd., Oulu, Finland) con abertura (*slit*) de 30 μm , y un lente OPT-000034 (SWIR, 16mm, f/1.4, Xenics, Leuven, Bélgica) con una distancia focal de 16 mm. Se utilizó una plataforma móvil de activación manual por sistema hidráulico para mover la muestra a velocidad constante a través de la zona de escaneo (campo de visión de la cámara).

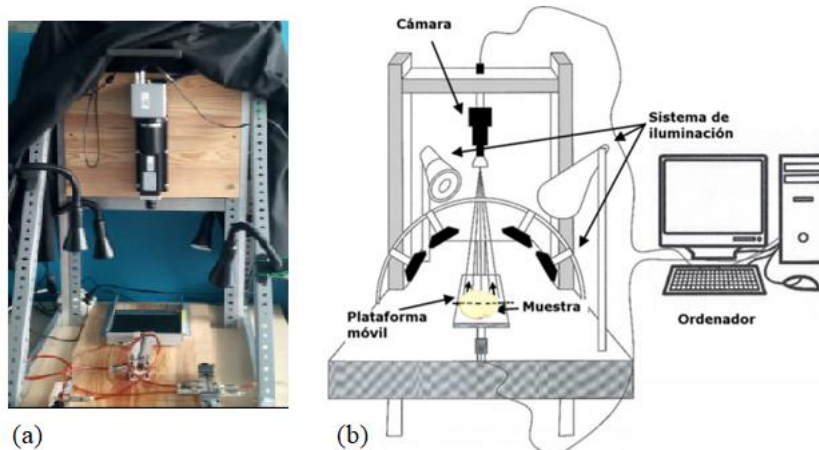


Figura 14. Equipo de adquisición de imágenes hiperespectrales: equipo real (a); esquema representativo (b) Fuente: adaptado de (Castro, 2017; López-Maestresalas et al., 2016)

Las muestras se escanearon a una velocidad de la plataforma de 9-10 mm/s y fijando una distancia de separación entre la lente y la superficie de la muestra de 30 cm. Además, se fijó una distancia de enfoque de 0,25 m y la máxima apertura del diafragma (f/1.4) en el lente. Las imágenes se adquirieron en el modo de calibración TrueNUC-HG de la cámara, a la máxima velocidad de escaneo (100 Hz), y con un tiempo de integración de 2000 μs . De esta manera, las imágenes tenían las siguientes dimensiones: 320 columnas (0,75 mm/pixel), 256 λ (cada 3,125 nm) y un número variable de filas (en función del tamaño de la rebanada de la papa, filas escaneadas).

Como fuente de iluminación de las muestras se utilizaron cuatro lámparas halógenas de 46 W (Lexman) que emiten radiación en el espectro infrarrojo. Fueron colocadas de tal manera que cada lámpara enfocaba una esquina de la placa porta muestras con el fin de conseguir una iluminación homogénea en el campo visual de la cámara. Para evitar interferencias por radiaciones externas de la luz ambiental, todo el sistema se cubrió con una manta negra durante el proceso de adquisición de imágenes y así evitar datos erróneos.

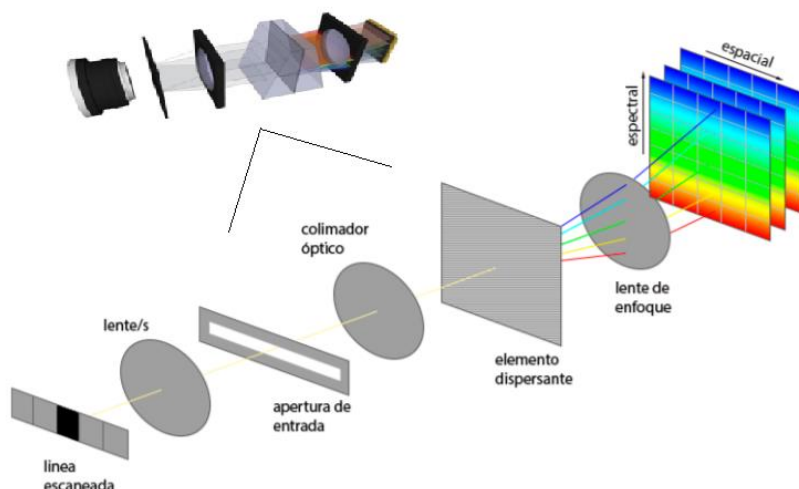


Figura 15. Esquema del funcionamiento del espectrógrafo en el sistema de adquisición hiperespectral. Fuente: adaptado de (Durán, 2011).

Para corregir la reflectancia relativa de las imágenes con el máximo y mínimo nivel de reflectancia que es capaz de captar el receptor, se realizó la calibración de la reflectancia capturando patrones de alto y bajo coeficiente de reflexión, llamados comúnmente “blanco” y “negro”.

3.4 Software

El sistema de adquisición hiperespectral se controla por medio de un ordenador equipado con el software Xeneth Advanced, el cual permite establecer los parámetros de ensayo de la cámara y controlar el funcionamiento de ésta durante la adquisición de las imágenes. Con MATLAB® se realizó el análisis y procesado de las imágenes hiperespectrales.

3.5 Procesado de imágenes

El proceso de captura generó matrices en formato PNG con 320 píxeles (columnas) y 256 (longitudes de onda), las imágenes se procesaron por medio del software MATLAB®. Para poder realizar la calibración y validación del modelo se construyó el hipercubo a partir de las imágenes obtenidas con el sistema HSI. Como resultado se obtuvo un hipercubo con determinado número de filas escaneadas más los píxeles y las longitudes de onda analizadas. Una vez obtenidos los hipercubos de cada muestra se realizó el procesado de las imágenes hiperespectrales.

- **Normalización**

Una vez obtenido el hipercubo, es necesario normalizar los valores de reflectancia obtenidos, ya que su rango varió de 0 a 65535 niveles de gris. La resolución radiométrica de este sistema son 16 bits, es decir, las imágenes adquiridas en escala de grises tienen 16 bits (o 2 Bytes) por píxel, lo que se traduce en 2^{16} (65536) niveles de gris, desde el 0 (negro) al 65535 (blanco). Para normalizar estos valores a reflectancia relativa se escalan todos los valores de reflectancia de los hipercubos a valores entre 0 y 1, usando como referencia dos patrones blanco y negro según la ecuación [1], descrita a continuación.

$$R_r = \frac{R_m - R_n}{R_b - R_n}; \quad [1]$$

Donde, R_r : Reflectancia relativa, R_m : Reflectancia de la muestra, R_n : Reflectancia mínima o negro y R_b : Reflectancia máxima o blanco.

- **Segmentación**

Antes de trabajar con los datos se dividió la información en 2 grupos (calibración y validación), para ello se utilizó un algoritmo del software MATLAB®. Se crearon los siguientes grupos: el 75% de la información se utilizó para la calibración y con esta información se elaboró el modelo de análisis discriminante y el 25% restante de la información se utilizó para la validación externa con la que se comprobó la eficacia del modelo para clasificar muestras.

Luego de normalizar los hipercubos se descartó aquella información que no tiene interés para el estudio. Para las imágenes utilizadas en calibración se manejó la herramienta HYPER-Tools dentro del entorno MATLAB. Para ello, se seleccionaron de forma manual 10 píxeles de cada una de las imágenes, de 27 imágenes de la clase de cocción y de 27 de la clase de fritura. De esta manera para cada muestra se tiene una matriz con dimensiones 10×224 . El vector de longitud de ondas se recortó de 256 a 224 para mejorar la calidad del estudio (eliminar ruido instrumental) con una dimensión de λ 1004-1700nm. En la figura 16, se observa el proceso de selección de los 10 píxeles considerando el área completa de la muestra, con el fin de conseguir una representación homogénea de los datos.

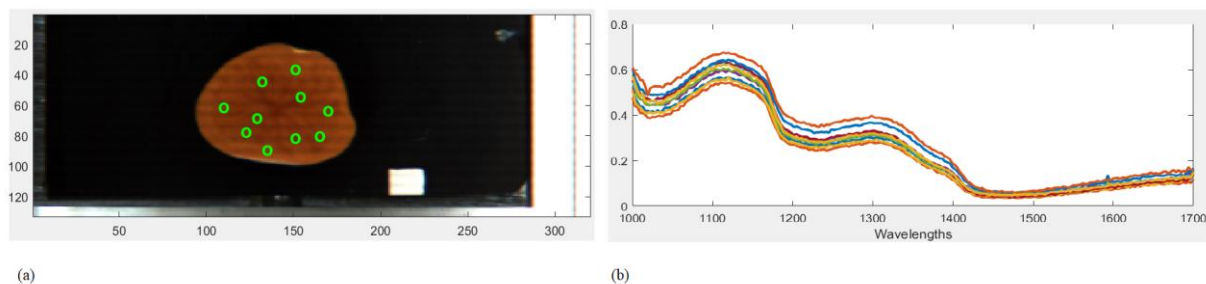


Figura 16. Selección de píxeles (a) y sus respectivos espectros (b). Fuente: elaboración propia.

Para la validación externa se utilizaron 18 imágenes, con las cuales se generó una matriz 2D para cada muestra y con el fin de separar el fondo (plataforma) de la muestra de papa se creó una máscara combinada con el hipercubo normalizado, como resultado se obtuvo una imagen binaria donde se asignó el valor 0 a los píxeles del fondo y el valor 1 a los de la muestra (ROIs). Al igual que la calibración se eliminaron del espectro las primeras 32 longitudes de onda. En esta investigación se utilizaron el 100% de los píxeles de cada muestra, con el objetivo de crear una imagen química para obtener un mapa representativo de cada una de ellas. La imagen química permitió diferenciar las muestras de papa según su aptitud de transformación industrial (frito o cocido).

3.6 Análisis quimiométrico: análisis multivariante

Con el fin de clasificar las papas en los grupos considerados (fritura y cocción) a través de sus espectros NIR, se utilizó la herramienta **PLS-Toolbox** bajo el entorno de MATLAB. Se aplicó el método cualitativo de análisis estadístico **PLS-DA** (*Partial Least Square Discriminant Analysis*). Este método combina la regresión lineal por mínimos cuadrados parciales con el análisis discriminante, mediante el cual se obtuvo la separación entre las clases. De esta forma, el modelo estableció una relación entre las variables predictivas X (valores de reflectancia en las longitudes de onda medias) y la variable dependiente Y (clases en los datos). El número de variables latentes a considerar durante el desarrollo del modelo se estableció en base al modelo que minimiza tanto el error medio de calibración como de validación cruzada. Por tanto, el modelo se eligió teniendo en cuenta la mínima distancia (los valores de errores), y el mayor porcentaje de varianza explicada

Se aplicaron los siguientes pretratamientos a los datos: suavizado (*Smoothing Savitzky-Golay*), un proceso de normalización ponderada (*Standar Normal Variate scaling, SNV*), primera y segunda derivada de los datos (*Derivate Savitzky-Golay*), un proceso de eliminación de los

valores medios (mean centering) y combinaciones de dichos pretratamientos. Todos los pretratamientos se realizaron con el fin de que el modelo creado fuera capaz de identificar en base a la información espectral analizada, la clase a la que pertenece cada papa (fritura o cocción).

- **Validación cruzada**

Este proceso se realizó para optimizar el modelo y garantizar la independencia de este dentro de la calibración. La validación cruzada, o cross-validation, se basa en extraer sucesivamente datos del conjunto de calibración y utilizarlos para validar el modelo. En este estudio se utilizó el método de cortinas venecianas (Venetian Blinds), con 10 divisiones (*splits*) y 1 muestra por división. Por tanto, se divide la muestra de entrenamiento en diferentes paquetes (folders), y mide el éxito de clasificación en cada uno tras realizar un entrenamiento con el resto.

Para evaluar la efectividad de los modelos, se utilizaron diferentes criterios: mediante la matriz de confusión, se obtuvo la **sensibilidad**, la **especificidad**, el **error de clase** y el **porcentaje de muestras correctamente clasificadas** en cada clase, así como la varianza explicada.

Matriz de confusión: es una matriz de $N \times N$ donde los valores de su diagonal principal serán aquellos predichos correctamente, es decir los verdaderos positivos y verdaderos negativos, mientras que los valores fuera de la diagonal corresponderán a los falsos positivos y falsos negativos:

- Verdadero positivo (VP): Indica la cantidad de muestras de la clase “A” clasificadas correctamente en la clase “A”. En este estudio, los verdaderos positivos serán aquellos píxeles de la clase cocción que sean correctamente clasificados como pertenecientes a esa clase cocción.
- Falso negativo (FN): Muestras de la clase “A” que han sido clasificadas como clase “B”, es decir, aquellos píxeles que, aun perteneciendo a la clase cocción, han sido clasificados como pertenecientes a la clase fritura.
- Verdadero negativo (VN): Indica la cantidad de muestras de la clase “B” clasificadas en la clase “B”. En este caso, serán aquellos píxeles pertenecientes a la clase fritura que han sido correctamente clasificados en su clase.

- Falso positivo (FP): Muestras de la clase “B” que han sido clasificadas como clase “A”., es decir, píxeles que perteneciendo a la clase fritura han sido incorrectamente clasificados como de clase cocción.

Con las directrices anteriores se definió la sensibilidad y la especificad:

Sensibilidad: con valores entre 0 y 1, da la proporción de los valores positivos que han sido correctamente clasificados como tal, por lo tanto, conforme disminuye de 1, el modelo tiene dificultades para diferenciar las muestras de cada grupo. Se obtiene con la ecuación [2].

$$\text{Sensibilidad} = \frac{VP}{VP+FN}; \quad [2]$$

Especificidad: con valores entre 0 y 1, da la proporción de verdaderos negativos (VN). Se obtiene con la ecuación [3].

$$\text{Especificidad} = \frac{VN}{FP+FN}; \quad [3]$$

Error de clase: es el promedio de las tasas de falsos positivos y falsos negativos para cada clase. Se calcula según la ecuación [4]

$$\text{Error} = 1 - \frac{(\text{Especificidad} + \text{sensibilidad})}{2}; \quad [4]$$

Validación externa: para realizar validación del modelo creado se utilizaron las matrices de las muestras de validación, las cuales no se han considerado en la elaboración del modelo de calibración.

4. RESULTADOS Y DISCUSIÓN

4.1 Modelos de calibración para la clasificación de la papa según su aptitud industrial

Se generaron 13 modelos (tabla 7) mediante el análisis PLS-DA con los datos destinados a la calibración, los cuales se validaron internamente mediante la validación cruzada. Luego se realizó una validación externa con cada uno de los modelos generados, tal como se describe en el inciso 3.6. Los modelos de calibración se desarrollaron con diferentes pretratamientos descritos en el mismo inciso. El modelo número 1 de calibración se realizó con los datos en “crudo”, sin ningún tipo de pretratamiento. Este modelo fue usado como referencia para contrastar los efectos de los distintos pretratamientos.

Tabla 7. Pretratamientos utilizados en los modelos desarrollados.

Modelo	Pretratamientos
1	Crudo
2	Smoothing
3	SNV
4	1ª Derivada
5	2ª Derivada
6	Mean Center
7	Smoothing +SNV+ 1ª Derivada
8	Smoothing +SNV+2ª Derivada
9	Smoothing +SNV+Mean Center
10	Smoothing +1ª Derivada +Mean Center
11	Smoothing +2ª Derivada +Mean Center
12	SNV+1ª Derivada +Mean Center
13	SNV+2ª Derivada +Mean Center

En la tabla 8 se muestra el número de muestras utilizado para la elaboración de cada modelo, el número de datos eliminados (*outliers*), el número de variables latentes (LV), la varianza que explica el modelo, así como el porcentaje de muestras clasificadas correctamente (CC) en la validación cruzada (CV). Para la selección del número de variables latentes se tuvo en cuenta el error estándar del modelo calculado y el de la validación cruzada de cada uno de los modelos. El número de muestras no consideradas para la construcción del modelo (N) supone una pérdida de datos inferior al 10% en todos los casos.

En la tabla 8 se han marcado los 3 modelos con mejores resultados, en base al porcentaje de CC (CV). Cabe destacar que, para su elección se han tenido en cuenta otros parámetros que no aparecen en la tabla y se detallan más adelante, como la sensibilidad, la especificidad y el error,

tanto de la calibración como de la validación cruzada. En ningún caso de estudio los modelos clasificaron las muestras en la categoría sin asignar, por lo que no se agregó en la tabla.

Tabla 8. Características de cada uno de los modelos.

Modelo	N Muestras	Outliers	Pretratamiento	% Varianza	LV	% CC (CV)	
						Cocción	Fritura
1	513	27	Crudo	99,98	5	63,20	71,86
2	525	15	Smoothing	100	6	72,69	70,94
3	516	24	SNV	99,97	4	67,55	69,29
4	516	24	1ª Derivada	99,86	5	69,38	70,15
5	520	20	2ª Derivada	99,18	6	68,58	66,79
6	540	0	Mean Center	98,97	4	70,74	68,51
7	516	24	SM+SNV+1ª D	99,94	4	68,34	69,65
8	535	5	SM+SNV+2ª D	99,76	4	69,66	67,91
9	502	38	SM+SNV+MC	97,38	5	76,30	68,77
10	494	46	SM+1ª D+MC	98,74	8	81,74	79,33
11	520	20	SM+2ª D+MC	92,04	9	70,88	66,02
12	535	5	SNV+1ª D+MC	96,40	7	70,37	71,37
13	535	5	SNV+2ª D+MC	87,13	6	69,92	65,42

Los modelos considerados como mejores son el número 2 (Smoothing), 9 (SM+SNV+MC) y 10 (SM+1ª D+MC). En los tres modelos se aplica el pretratamiento Smoothing, en dos de los casos es acompañado por Mean Center. En la figura 17 observamos el efecto de los pretratamientos sobre los espectros en crudo. En los espectros sin pretratamiento se observa de forma generalizada un cierto ruido espectral y dispersión de los datos. Con el pretratamiento Smoothing se consiguió un suavizado de los datos, corrigiendo el ruido y la dispersión. Con los pretratamientos SM + SNV + MC se normalizaron los datos sin embargo se observa ruido en los mismos. Con los pretratamientos SM+1ª D+MC se observa disminución del ruido espectral e información irregular. El orden en el que se aplican los pretratamientos al modelo influye en los resultados, por ello en esta investigación se han aplicado de la manera que mejores resultados se obtienen.

En la tabla 8 observamos que los porcentajes de muestras clasificadas correctamente en la validación cruzada varían dependiendo de los pretratamientos aplicados, obteniendo los mejores resultados el modelo 10 (SM + 1ª D + MC) con un porcentaje de muestras clasificadas correctamente de 81,74% para la aptitud de cocción y 79,33% para la aptitud de fritura y con 8 variables latentes. Los investigadores (Elmasry et al., 2012) consideran un porcentaje de muestras correctamente clasificadas $\geq 75\%$ razonable, tal como lo mencionan en su estudio de clasificación en línea de papas irregulares mediante el uso de imágenes hiperespectrales. En base a estos resultados se puede considerar que el modelo es capaz de diferenciar entre ambas clases.

En la tabla 9 se observa que la especificidad, sensibilidad y error de clase tanto de la calibración como de la validación cruzada varían con los pretratamientos aplicados. De acuerdo con esta tabla se confirma que el modelo 10 (SM + 1ª D + MC) clasifica bien las clases. Este es el modelo que presenta los mejores resultados en sensibilidad, especificidad (más cerca de 1) y el error más bajo, tanto en calibración como en validación cruzada. Dado que los valores de sensibilidad y especificidad son similares, el error de clase está equilibrado, es decir, no hay una tendencia particular en el modelo a clasificar píxeles de la clase fritura como cocción o *viceversa*. Con los modelos 2 (Smoothing) y 9 (SM+SNV+MC) se obtuvieron resultados de especificidad menores a los resultados obtenidos con el modelo sin pretratamiento, tanto en la calibración como en la validación. Estos resultados afectan directamente el porcentaje de muestras clasificadas correctamente en la validación cruzada, específicamente en las variedades de papa con aptitud de fritura (tabla 8). En ningún caso de estudio los errores fueron mayores al error del modelo sin pretratamiento. Para la continuidad de este estudio se descartaron estos dos (2 y 9) modelos.

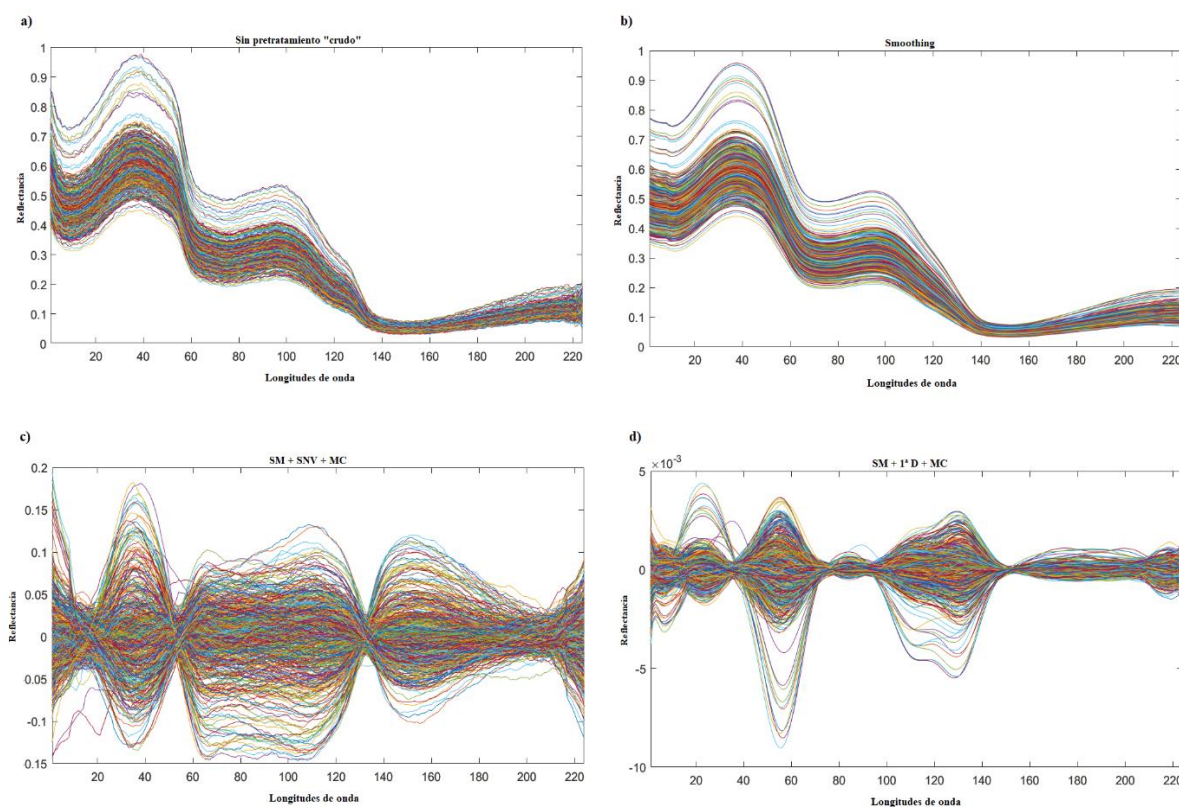


Figura 17. Efecto de los pretratamientos en los espectros de reflectancia: a) Sin pretratamiento, b) Smoothing, c) SM + SNV + MC y d) SM + 1ª D + MC.

Tabla 9. Resultados de la sensibilidad y especificidad de los 3 mejores modelos y los datos sin pretratamiento.

Modelo	Pretratamiento	Sensibilidad (Cal)	Especificidad (Cal)	Error (Cal)	Sensibilidad (CV)	Especificidad (CV)	Error (CV)
1	Crudo	0,636	0,734	0,315	0,632	0,719	0,324
2	Smoothing	0,742	0,717	0,270	0,735	0,713	0,276
9	SM+SNV+MC	0,775	0,692	0,266	0,763	0,688	0,274
10	SM+1D+MC	0,833	0,810	0,178	0,817	0,793	0,194

4.2 Validación externa del modelo de calibración

Una vez generado el modelo de calibración se realiza la validación externa del mismo. Para ello se generó una matriz de validación con el 25% de los datos. Es decir, la validación externa se realizó de manera conjunta sobre el total de las muestras destinadas para ese fin. La matriz de validación contiene información de 18 muestras, 9 variedades con aptitud de cocción y 9 con aptitud de fritura. Para ello se ha necesitado obtener de cada muestra su respectivo hipercubo normalizado (eliminando las 32 primeras longitudes de onda) y la máscara. Las muestras utilizadas en la validación externa no fueron utilizadas para crear el modelo de calibración, tal y como se describe en el inciso 3.6 de este estudio.

En la tabla 10 se muestran los resultados obtenidos en la validación externa del modelo 10 (SM + 1ª D + MC). Se observa que el porcentaje de muestras correctamente clasificadas para la aptitud de cocción es de 66,44% y para fritura es 64,38%. El error de predicción del modelo fue de 0,35. Aunque los porcentajes de predicción son relativamente bajos y el error es alto, son datos de un primer estudio que posibilitan la clasificación de las muestras utilizando únicamente la información espectral. Además, hay que tener en cuenta que estos resultados se generaron en base a una matriz que contiene el total de las muestras de validación. En la figura 18, se observa gráficamente los porcentajes de predicción que ejecuta el modelo sobre la matriz de calibración externa. Las muestras por arriba de la línea negra representan el porcentaje de las muestras bien clasificadas en su grupo y por debajo las mal clasificadas. La figura 18 representa la clasificación externa de las muestras pertenecientes a la clase cocido. En este caso, lo ideal sería encontrar todas las muestras de clase cocido (diamante rojo) por encima de la línea de discriminación y todas las muestras de la clase fritura (cuadrado verde) por debajo de ésta. Sin embargo, son pocos los modelos de este tipo que logran un 100% de clasificación correcta.

Los investigadores (López-Maestresalas et al., 2016) reportaron en su estudio un porcentaje de muestras correctamente clasificada superior al 90%. Sin embargo, las clases presentan una diferenciación que puede llegar a ser evidente incluso dentro del espectro visible, como por ejemplo en detección de defectos en papa. Esta característica difiere de este estudio ya que las clases de papa no son evidentes dentro del espectro visible. Muchos autores realizan una caracterización de las clases en base a análisis cuantitativos con una exploración de los datos o una selección de las principales longitudes de onda que mejor representen las clases de discriminación (Ivorra, 2015).

Tabla 10. Resultados de la validación externa del modelo de calibración.

Modelo	Preprocesado	Sensibilidad (Pred)	Especificidad (Pred)	Error (Pred)	% Bien Clasificadas	
					Cocción	Fritura
10	SM+1D+MC	0,664	0,644	0,345	66,44	64,38

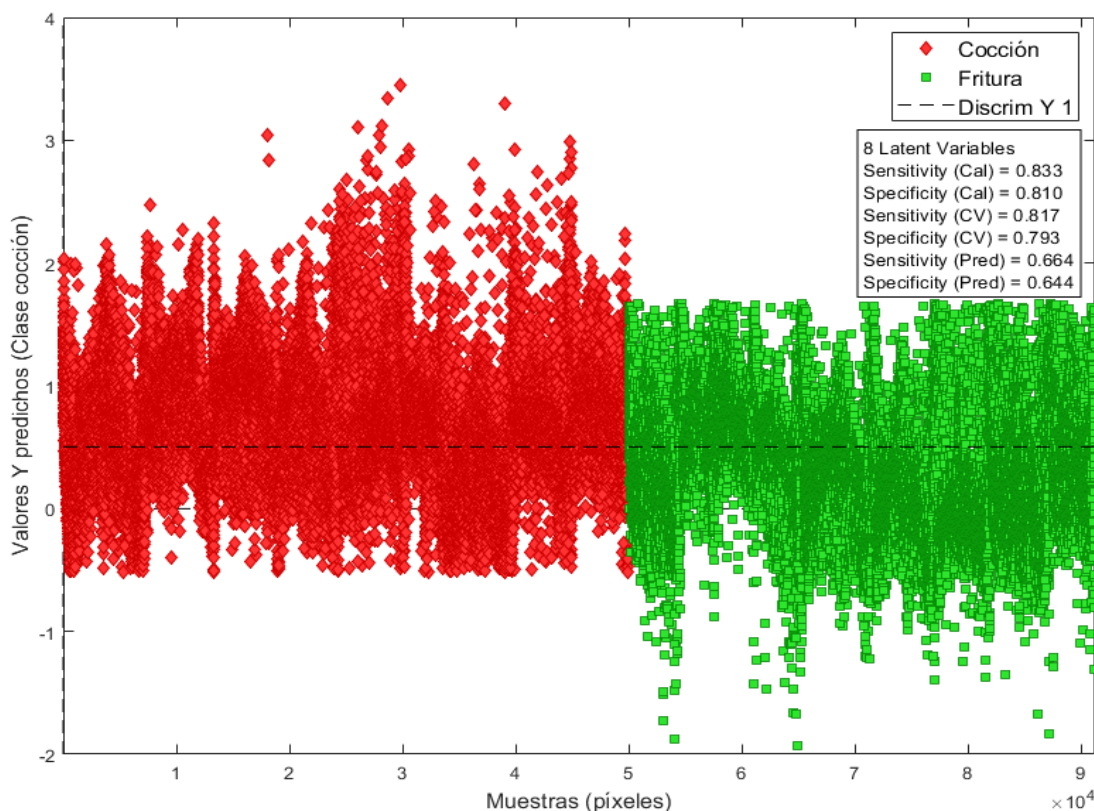


Figura 18. Representación gráfica de los resultados de predicción que ejecuta el modelo sobre la matriz de calibración externa.

En la figura 19 se presenta la imagen química de la muestra de validación externa. Esta imagen se realizó con la información espectral obtenida de la composición de cada muestra de papa. En general se observa que la mayor parte de los píxeles son correctamente clasificados. Esto indica que el modelo podría clasificar de manera correcta una muestra en los diferentes grupos, aunque los porcentajes de bien clasificadas sean 66 y 64%. Algunas muestras, como la 1ª y la 7ª de la clase de cocción tienen muchos FN de igual forma la 2ª y la 3ª de la clase de fritura presentan muchos FP. Esto se podría deber al error de clase del modelo por la similitud en la información espectral.

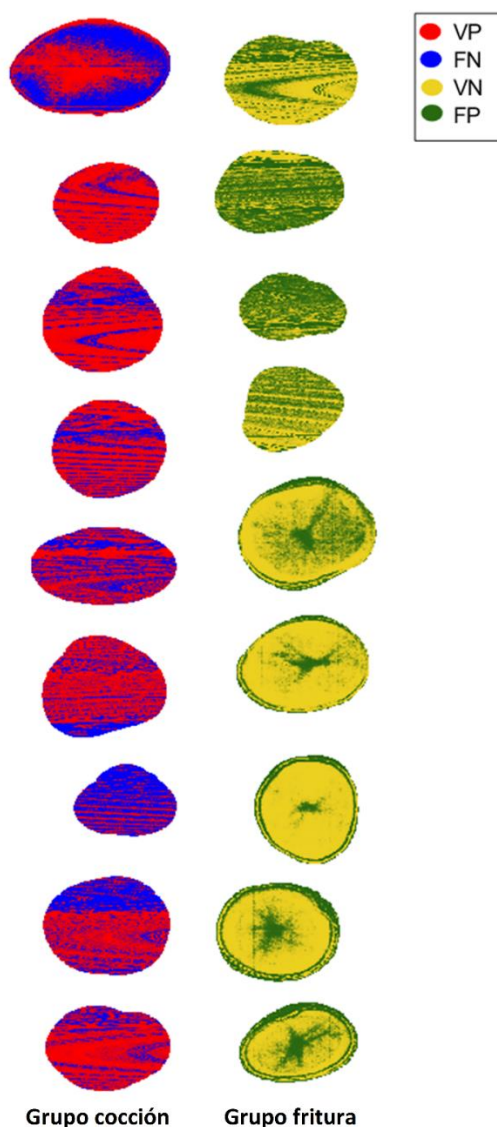


Figura 19. Imagen química de las muestras utilizadas en la validación externa.

5. CONCLUSIONES

A partir de los resultados obtenidos en este estudio, se concluye lo siguiente:

- Los mejores resultados de calibración se obtuvieron con tres modelos en los que se ha utilizado el pretratamiento Smoothing, consiguiendo un porcentaje de muestras correctamente clasificadas superior al 68%.
- El modelo de calibración con el cual se obtuvo un porcentaje de muestras correctamente clasificadas de 79,33% para el grupo de fritura y de 81,74% para el grupo de cocción, fue una combinación de los pretratamientos: Smoothing, 1ª Derivada y Mean Center.
- En la validación externa del modelo de clasificación se ha obtenido un porcentaje de muestras correctamente clasificadas superior al 64% según su aptitud industrial de fritura y cocción utilizando solo la información espectral de las muestras.
- Por medio de la información de las imágenes hiperespectrales se obtuvo la imagen química de la muestra de validación externa, donde se corrobora que en general el modelo desarrollado clasifica correctamente los píxeles en cada grupo. Sin embargo, algunas imágenes presentan píxeles incorrectamente clasificados. Debido a esto para futuros estudios se recomienda realizar una exploración de los datos o una selección de las principales longitudes de onda que mejor representen los grupos, con el objetivo de garantizar mejores resultados.

6. BIBLIOGRAFÍA

- AGSI/FAO. (2001). Potato: Post-harvest Operations. Post/harvest Operations, 2. Recuperado de http://www.fao.org/fileadmin/user_upload/inpho/docs/Post_Harvest_Compendium_-_Potato.pdf
- Andrade, H. (1997). Requerimientos cualitativos para la industrialización de papa. Revista Informativa INIAP (9):21-23, 4 p. https://doi.org/*EC-INIAP-BEESC-MGC. Quito (INIAP/Rev (9):21-23. 1997)
- Arce, F. A. (2016). Cultivo de Patata en Castilla y León. Cajamar Caja Rural, 1, 29-60. Recuperado de <https://www.publicacionescajamar.es/uploads/cultivos-hortícolas-al-aire-libre/01-cultivos-hortícolas-al-aire-libre.pdf>
- Atherton, D., Watson, D. G., Zhang, M., Qin, Z., & Liu, X. (2015). Hyperspectral Spectroscopy for Detection of Early Blight (*Alternaria solani*) Disease in Potato (*Solanum tuberosum*) Plants at Two Different Growth Stages. American Society of Agricultural and Biological Engineers, 1(3), 189-194. <https://doi.org/10.13031/aim.20152186976>
- Ayala, D. (2018). Automatización del análisis de imágenes hiperespectrales para identificación de aptitud de patatas. Universidad Pública de Navarra.
- Bernabé, E. S. (2016). Procesado de imágenes hiperespectrales. Escuela Técnica Superior De Ingeniería Y Sistemas De Telecomunicación, 65-91. Recuperado de http://oa.upm.es/44860/3/TFG_ESTHER_SANCHEZ_BERNABE.pdf<http://oa.upm.es/44860/>
- Birch, P. R. J., Bryan, G., Fenton, B., Gilroy, E. M., Hein, I., Jones, J. T., ... Toth, I. K. (2012). Crops that feed the world 8: Potato: Are the trends of increased global production sustainable? En Food Security (Vol. 4). <https://doi.org/10.1007/s12571-012-0220-1>
- Borruey A., Cotrina F., Mula J., V. C. (1999). Calidad industrial y culinaria de las variedades de patata. Congreso Iberoamericano de Investigación y Desarrollo en Patata.
- Castro, E. (2017). Aplicación de Imágenes Hiperespectrales para la Clasificación de Tubérculos de Patata de Acuerdo a su Aptitud para la Transformación Industrial. Universidad Pública de Navarra.

- Castro, W. (2015). *Aplicación De La Tecnología De Imágenes Hiperespectrales Al Control De Calidad De Productos Agroalimentarios De La Región De Amazonas (Perú)*. Universidad Politécnica de Valencia.
- Chang, C. I., & Du, Q. (2004). Estimation of number of spectrally distinct signal sources in hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 42(3), 608-619. <https://doi.org/10.1109/TGRS.2003.819189>
- Cho, B. K., Kim, M. S., Baek, I. S., Kim, D. Y., Lee, W. H., Kim, J., ... Kim, Y. S. (2013). Detection of cuticle defects on cherry tomatoes using hyperspectral fluorescence imagery. *Postharvest Biology and Technology*, 76, 40-49. <https://doi.org/10.1016/j.postharvbio.2012.09.002>
- CIP. (2012). *Nuevas variedades de papa (p. 55 pp)*. p. 55 pp. Recuperado de <http://cipotato.org/wp-content/uploads/2013/08/005909.pdf>
- CIP. (2018). *Centro Internacional de la Papa (p. 2)*. p. 2. Recuperado de <https://cipotato.org/es/about/estrategia-y-plan-corporativo-2014-23/>
- Condori, B., Casa, A. De, Soratto, P., Olarte, S., Mompies, J., Clavijo, N. L., ... García, C. (2016). *Modelación de la papa en Latinoamérica (JRC Scienc; European Union, Ed.)*. <https://doi.org/10.2788/11877>
- Dacal-Nieto, A., Formella, A., Carrión, P., Vazquez-Fernandez, E., & Fernández-Delgado, M. (2011a). Common scab detection on potatoes using an infrared hyperspectral imaging system. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6979 LNCS(PART 2), 303-312. https://doi.org/10.1007/978-3-642-24088-1_32
- Dacal-Nieto, A., Formella, A., Carrión, P., Vazquez-Fernandez, E., & Fernández-Delgado, M. (2011b). Non-destructive detection of hollow heart in potatoes using hyperspectral imaging. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6855 LNCS(PART 2), 180-187. https://doi.org/10.1007/978-3-642-23678-5_20
- Devaux, A., Kromann, P., & Ortiz, O. (2014). Potatoes for Sustainable Global Food Security. *Potato Research*, 57(3-4), 185-199. <https://doi.org/10.1007/s11540-014-9265-1>

Dorrepaal, R., Malegori, C., & Gowen, A. (2016). Tutorial: Time series hyperspectral image analysis.

Journal of Near Infrared Spectroscopy, 24(2), 89-107. <https://doi.org/10.1255/jnirs.1208>

Durán, D. (2011). Técnicas de oversampling aplicadas al análisis de imágenes hiperespectrales. Universidad de Extremadura.

El-Mesery, H. S., Mao, H., & Abomohra, A. E. F. (2019, febrero 2). Applications of non-destructive technologies for agricultural and food products quality inspection. Sensors (Switzerland), Vol. 19. <https://doi.org/10.3390/s19040846>

ElMasry, G., Cubero, S., Moltó, E., & Blasco, J. (2012). In-line sorting of irregular potatoes by using automated computer-based machine vision system. Journal of Food Engineering, 112(1-2), 60-68. <https://doi.org/10.1016/j.jfoodeng.2012.03.027>

ElMasry, G., Kamruzzaman, M., Sun, D. W., & Allen, P. (2012). Principles and Applications of Hyperspectral Imaging in Quality Evaluation of Agro-Food Products: A Review. Critical Reviews in Food Science and Nutrition, 52(11), 999-1023. <https://doi.org/10.1080/10408398.2010.543495>

ElMasry, G., Sun, D. W., & Allen, P. (2013). Chemical-free assessment and mapping of major constituents in beef using hyperspectral imaging. Journal of Food Engineering, 117(2), 235-246. <https://doi.org/10.1016/j.jfoodeng.2013.02.016>

ElMasry, G., Wang, N., ElSayed, A., & Ngadi, M. (2007). Hyperspectral imaging for nondestructive determination of some quality attributes for strawberry. Journal of Food Engineering, 81(1), 98-107. <https://doi.org/10.1016/j.jfoodeng.2006.10.016>

Esquivel, P., Brenes, A., & Alfaro, I. (2009). Propiedades Físico-Químicas Y Parámetros De Calidad Para Uso Industrial De Cuatro Variedades De Papa. Agronomía Costarricense, 33(1), 77-89.

FAO. (2006). Tesoro enterrado: La papa.

FAO. (2017). The state of agriculture and food. En Leveraging food systems for inclusive rural transformation. <https://doi.org/0251-1371>

FAO. (2019). FAOSTAT. Recuperado 2 de enero de 2020, de <http://www.fao.org/faostat/es/#data/QC>

Gallego, E. M., Miguez, M., & De la Montaña, J. (2006). Aptitud De Variedades De Patata Para Su Transformación En Barritas Prefritas Aptitude of Varieties of Potato for Its Industrial

Transformation in Fried Potatoes Slices. *Ciencia y Tecnología Alimentaria*, 5(3), 189-194.
<https://doi.org/10.1080/11358120609487691>

García-Segovia, P., Andrés-Bello, A., & Martínez-Monzó, J. (2008). Textural properties of potatoes (*Solanum tuberosum* L., cv. Monalisa) as affected by different cooking processes. *Journal of Food Engineering*, 88(1), 28-35. <https://doi.org/10.1016/j.jfoodeng.2007.12.001>

Gardi, C., Angelini, M., Barceló, S., Comerma, J., Cruz Gaistardo, C., Encina Rojas, A., Jones, A., Krasilnikov, P., M. S. B., M.L., Montanarella, L., Muniz Ugarte, O., Schad, P., V. R., & M.I., Vargas, R. (2014). Atlas de Suelos de América Latina y el Caribe (Luxembourg; Unión Europea, Ed.). <https://doi.org/10.2788/37334>

Garzón, F. J. M. (2007). Sociedades precolombinas asociadas a la domesticación y cultivo de la papa (*Solanum tuberosum*) en Sudamérica Pre-Columbian societies associated with the domestication and cultivation of potato (*Solanum tuberosum*) in South America. 14, 1-9.

Gowen, A. A., O'Donnell, C. P., Cullen, P. J., Downey, G., & Frias, J. M. (2007). Hyperspectral imaging - an emerging process analytical tool for food quality and safety control. *Trends in Food Science and Technology*, 18(12), 590-598. <https://doi.org/10.1016/j.tifs.2007.06.001>

Grahn, H. F., & Geladi, P. (2007). Techniques and Applications of Hyperspectral Image Analysis. *Techniques and Applications of Hyperspectral Image Analysis*, 1-368. <https://doi.org/10.1002/9780470010884>

Granados, C., Acevedo, D., Cabeza, A., & Lozano, A. (2014). Análisis de perfil de textura en plátanos Pelipita, Hartón y Topocho. *Informacion Tecnologica*, 25(5), 35-40. <https://doi.org/10.4067/S0718-07642014000500006>

H.Bohl, W., B.Johnson, S., Belyea, S., Brown, C., Bushway, A., & Bussan, A. (2010). Commercial potato production in North America. *American Journal of Potato Research*, 87 suppl, 1-90.

INIA. (2017). Manual Del Cultivo De La Papa En Chile. Inia, 144. Recuperado de <http://www.inia.cl/wp-content/uploads/ManualesdeProduccion/10 Manual Papa.pdf>

Ivorra, E. (2015). Desarrollo de técnicas de visión hiperespectral y tridimensional para el sector agroalimentario (Universidad Politécnica de Valencia). Recuperado de <https://riunet.upv.es/bitstream/handle/10251/48541/Ivorra - Desarrollo de técnicas de visión hiperespectral y tridimensional para el sector agroali....pdf?sequence=1>

- Jansky, S. H., Jin, L. P., Xie, K. Y., Xie, C. H., & Spooner, D. M. (2009). Potato production and breeding in China. *Potato Research*, 52(1), 57-65. <https://doi.org/10.1007/s11540-008-9121-2>
- Ji, Y., Sun, L., Li, Y., & Ye, D. (2019). Detection of bruised potatoes using hyperspectral imaging technique based on discrete wavelet transform. *Infrared Physics and Technology*, 103(September), 103054. <https://doi.org/10.1016/j.infrared.2019.103054>
- Juszczak, L., Witczak, M., Zieba, T., & Fortuna, T. (2012). Rheological behaviour of heated potato starch dispersions. *International Agrophysics*, 26(4), 381-386. <https://doi.org/10.2478/v10247-012-0053-3>
- Keijbets, M. J. H. (2008). Potato processing for the consumer: Developments and future challenges. *Potato Research*, 51(3-4), 271-281. <https://doi.org/10.1007/s11540-008-9104-3>
- Koch, M., Naumann, M., Pawelzik, E., Gransee, A., & Thiel, H. (2020). The Importance of Nutrient Management for Potato Production Part I: Plant Nutrition and Yield. *Potato Research*, 63(1), 97-119. <https://doi.org/10.1007/s11540-019-09431-2>
- Lammertyn, J., Peirs, A., De Baerdemaeker, J., & Nicolai, B. (2000). Light penetration properties of NIR radiation in fruit with respect to non-destructive quality assessment. *Postharvest Biology and Technology*, 18(2), 121-132. [https://doi.org/10.1016/S0925-5214\(99\)00071-X](https://doi.org/10.1016/S0925-5214(99)00071-X)
- López-Maestresalas, A. (2016). Near-infrared Spectroscopy and Hyperspectral Imaging for Non-destructive Quality Inspection of Potatoes. Universidad Pública de Navarra.
- López-Maestresalas, A., Keresztes, J. C., Goodarzi, M., Arazuri, S., Jarén, C., & Saeys, W. (2016). Non-destructive detection of blackspot in potatoes by Vis-NIR and SWIR hyperspectral imaging. *Food Control*, 70, 229-241. <https://doi.org/10.1016/j.foodcont.2016.06.001>
- Loyola, N., Oyarce, E., & Acuña, C. (2010). Evaluación Del Contenido De Almidón En Papas (N.º Carmen 684). Curicó - Chile.
- Luján Claire, L. (1996). Tomado de la Revista Papa, Órgano Informativo de la Federación Colombiana de Productores de Papa.
- MAGRAMA. (2018). Superficie Producción Valor (miles de hectáreas) (miles de toneladas). Recuperado 8 de marzo de 2020, de MAGRAMA website: <https://www.mapa.gob.es/estadistica/pags/anuario/2018/CAPITULOSPDF/CAPITULO07/pdf>

c07_3.2.3.pdf

Maza, S. (2011). La papa nuestra de cada día (p. 13). p. 13. Lima.

Ministerio de Agricultura y Pesca, A. (2018). Informe del Consumo Alimentario en España 2018. Gobierno de España, 538. <https://doi.org/10.1002/chin.200343156>

Mollazade, K., Omid, M., Tab, F. A., & Mohtasebi, S. S. (2012). Principles and Applications of Light Backscattering Imaging in Quality Evaluation of Agro-food Products: A Review. *Food and Bioprocess Technology*, 5(5), 1465-1485. <https://doi.org/10.1007/s11947-012-0821-x>

Muñoz, M. (2014). Composición y aportes nutricionales de la papa. *Revista Agrícola*, 36-37. Recuperado de http://www.inia.cl/wp-content/uploads/2014/09/revista_agricola_octubre_36-37.pdf

Nicolaï, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I., & Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology*, 46(2), 99-118. <https://doi.org/10.1016/j.postharvbio.2007.06.024>

Ozaki, Y., Huck, C. W., & Beć, K. B. (2018). Near-IR Spectroscopy and Its Applications. En V. P. Gupta (Ed.), *Molecular and Laser Spectroscopy: Advances and Applications* (pp. 11-38). <https://doi.org/10.1016/B978-0-12-849883-5.00002-4>

Pasquini, C. (2018). Near infrared spectroscopy: A mature analytical technique with new perspectives – A review. *Analytica Chimica Acta*, 1026, 8-36. <https://doi.org/10.1016/j.aca.2018.04.004>

Prada, R. (2008). Alternativa de Aprovechamiento Eficiente de Residuos Biodegradables: el caso del almidón residual derivado de la industrialización de la papa. 182-192.

Román, M., & Hurtado, G. (2002). Cultivo de la Papa. *Centro Nacional De Tecnología Agropecuaria y Forestal*, 34.

Rosenfeld, A., & Kak, A. C. (1982). Digital Picture Processing, Volume 1. En W. Rheinboldt (Ed.), *Journal of Chemical Information and Modeling* (Segunda Ed, Vol. 53). <https://doi.org/10.1017/CBO9781107415324.004>

Rubén, F., José, M., André, D., Claudio, V., Alejandro, M., Daniela, C., & Esteban, L. (2017). VII Congreso ecuatoriano de la papa. En P. Kromann, C. Xavier, M. Byron, C. Patricio, L.-R.

Antonio, & C. Andrés (Eds.), Centro Internacional de la Papa – CIP. (p. 157). Quito: Centro Internacional de la Papa – CIP.

Rutherford, C. I. P. H. (2014). Strategy and corporate plan 2014–23 ©. Recuperado de <https://cipotato.org/es/about/estrategia-y-plan-corporativo-2014-23/>

Señán, G. B. (2010). El consumo en tiempos de crisis: una aproximación sociológica a la distribución del gasto en España. *Revista de Ciencias Sociales*, 45, 1-24.

Shimizu, T. and G. J. S. R. (2014). Los Supermercados y Cambios en la Cadena Productiva para la Papa en el Perú.

Su, Q., Kondo, N., Li, M., Sun, H., & Al Riza, D. F. (2017). Potato feature prediction based on machine vision and 3D model rebuilding. *Computers and Electronics in Agriculture*, 137, 41-51. <https://doi.org/10.1016/j.compag.2017.03.020>

Sun, D.-W. (2010). Hyperspectral Imaging for Food Quality Analysis and Control. En D.-W. Sun (Ed.), *Open Nutraceuticals Journal* (Vol. 4). <https://doi.org/10.2174/1876396001104010025>

Taylor, M. A., McDougall, G. J., & Stewart, D. (2007). Potato flavour and texture. En *Potato Biology and Biotechnology: Advances and Perspectives*. <https://doi.org/10.1016/B978-044451018-1/50066-X>

USIL. (2019). *La Papa: Orgullo del Perú* - Fondo Editorial USIL - Google Libros. Recuperado de [https://books.google.es/books?id=yDCNDwAAQBAJ&pg=PT268&lpg=PT268&dq=que+los+mayores+niveles+encontrados+de+acrilamida+están+en+alimentos+amiláceos+\(patatas+y+cereales\)&source=bl&ots=cda73HOYK4&sig=ACfU3U0DEwWE2F7_T5kJCbI7GPZY00OD5w&hl=es-419&sa=X&ved=2ahU](https://books.google.es/books?id=yDCNDwAAQBAJ&pg=PT268&lpg=PT268&dq=que+los+mayores+niveles+encontrados+de+acrilamida+están+en+alimentos+amiláceos+(patatas+y+cereales)&source=bl&ots=cda73HOYK4&sig=ACfU3U0DEwWE2F7_T5kJCbI7GPZY00OD5w&hl=es-419&sa=X&ved=2ahU)

Valdunciel, J. M. (2008). Método de valoración para las patatas. INIA (Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria), 13. Recuperado de <http://wwwsp.inia.es/Investigacion/OtrasUni/DTEVPF/Unidades/CentrosEnsayo/EstacionEnsayos/Documents/Métodospatata.pdf>

Wenhao, S., Guishan, L., Jianguo, H., Songlei, W., Xiaoguang, H., Wei, W., & Longguo, W. (2014). Detection of external defects on potatoes by hyperspectral imaging technology and image processing method. *Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering*, 28(21), 221-228. <https://doi.org/10.3969/j.issn.1002->

6819.2012.21.031

- Wu, D., & Sun, D. W. (2013). Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review - Part I: Fundamentals. *Innovative Food Science and Emerging Technologies*, 19, 1-14. <https://doi.org/10.1016/j.ifset.2013.04.014>
- Xiong, Z., Sun, D. W., Zeng, X. A., & Xie, A. (2014). Recent developments of hyperspectral imaging systems and their applications in detecting quality attributes of red meats: A review. *Journal of Food Engineering*, 132, 1-13. <https://doi.org/10.1016/j.jfoodeng.2014.02.004>