



UNIVERSIDAD
DE LA REPUBLICA
URUGUAY



Técnicas de procesamiento de imágenes aplicadas al monitoreo de procesos alimentarios

Mariana Estefani González Morales

Programa de Posgrado en Ingeniería Química
Facultad de Ingeniería
Universidad de la República

Montevideo – Uruguay
Marzo de 2021



UNIVERSIDAD
DE LA REPUBLICA
URUGUAY



Técnicas de procesamiento de imágenes aplicadas al monitoreo de procesos alimentarios

Mariana Estefani González Morales

Tesis de Maestría presentada al Programa de Posgrado en Ingeniería Química, Facultad de Ingeniería de la Universidad de la República, como parte de los requisitos necesarios para la obtención del título de Magister en Ingeniería Química.

Director de tesis:

Dra. Prof. Eliana Budelli Di Paola

Codirector:

Dr. Prof. Nicolás Leonardo Pérez Álvarez

Director académico:

Dra. Prof. Eliana Budelli Di Paola

Montevideo – Uruguay

Marzo de 2021

González Morales, Mariana Estefani

Técnicas de procesamiento de imágenes aplicadas al monitoreo de procesos alimentarios / Mariana Estefani González Morales. - Montevideo: Universidad de la República, Facultad de Ingeniería, 2021.

XIX, 99 p.: il.; 29, 7cm.

Director de tesis:

Eliana Budelli Di Paola

Codirector:

Nicolás Leonardo Pérez Álvarez

Director académico:

Eliana Budelli Di Paola

Tesis de Maestría – Universidad de la República, Programa de Ingeniería Química, 2021.

Referencias bibliográficas: p. 91 – 99.

1. Procesamiento de imágenes, 2. Monitoreo de procesos, 3. Quesos formadores de ojos, 4. Frutillas, 5. Atributos de calidad. I. Budelli Di Paola, Eliana et al. II. Universidad de la República, Programa de Posgrado en Ingeniería Química. III. Título.

INTEGRANTES DEL TRIBUNAL DE DEFENSA DE TESIS

Dra. Prof. Valeria Larnaudie

Dr. Prof. Federico Lecumberry

Dr. Prof. Ismael Núñez

Montevideo – Uruguay

Marzo de 2021

Dedicado a mis padres Ana y
Asdrúbal, y a mi hermano
Damián

Agradecimientos

Hay tantas personas y organizaciones que me apoyaron durante esta etapa a los cuales quisiera agradecer. En primer lugar, a mis tutores, Eliana y Nicolás, quienes se embarcaron en esta aventura de explorar el mundo del procesamiento de imágenes conmigo, que alentaron en mí la creatividad para la investigación y dedicaron su tiempo y energía a este proyecto. Les agradezco sus consejos, junto a su paciencia, comprensión, aliento y experiencia cuando más lo necesitaba.

Otra persona a quien quisiera agradecer es a Patricia Lema, quien me alentó a realizar la maestría y quien siempre estuvo presente con sus consejos desde que inicié mis primeras experiencias en el mundo de la investigación. Quisiera además agradecer al resto de los integrantes del grupo “Tecnologías Aplicadas a Procesos Alimentarios”: Sofía, Mariana, Florencia, Erika, Sylvia, Patricia y Magdalena. Es un placer poder trabajar con ustedes y compartir tantos momentos juntos. Este grupo es como una segunda familia para mí y estoy eternamente agradecida por las oportunidades que me han brindado.

También quiero dar las gracias a los distintos grupos del “Instituto de Ingeniería Química”, los cuales en reiteradas oportunidades me prestaron sus instalaciones y equipos cuando los necesitaba. A los integrantes del “Instituto de Ingeniería Química” por hacer que el ambiente laboral sea mucho más llevadero y por dedicarme su tiempo cuando necesitaba asistencia con algún equipo o experimento. En esta línea, quisiera agradecer particularmente a Viviana Palombo, Mariana Corengia y Jimena Ferreira, quienes fueron mis consejeras en momentos de frustraciones y dudas que surgen en la vida.

En cuanto a las instituciones que me permitieron acceder a los recursos para que esta maestría fuese efectuada, quisiera agradecer a la “Comisión Académi-

ca de Posgrado” por la beca económica que permitió que llevara adelante estos estudios, a la Facultad de Ingeniería (UdelaR) por el uso de instalaciones y los cursos necesarios, a la “Comisión Sectorial de Investigación Científica” por la financiación de proyectos de investigación y asistencia a congresos, y al “Espacio Intedisciplinario” de la Universidad de la República por la financiación de proyectos de investigación.

Por último pero no menos importantes, quisiera agradecer a mi familia y amigos, quienes me apoyaron en esta etapa y supieron escucharme tanto en los momentos más emocionantes como los de mayor incertidumbre que ocurrieron a lo largo de esta maestría.

Gracias totales.

RESUMEN

El aspecto visual de los alimentos es uno de los más relevantes a la hora de escoger un producto, particularmente para los productos con gran variabilidad tales como hortalizas, productos cárnicos y derivados lácteos, entre otros. Desviaciones del aspecto esperado del alimento pueden ser asociadas a un producto inferior por parte de los consumidores, impactando negativamente en la comercialización de dicho producto. Un método para categorizar los productos alimenticios en distintos niveles de calidad basándose en su aspecto es la inspección tradicional, tarea que involucra la participación directa de personas en el proceso de evaluación. Sin embargo, este método es una tarea tediosa, subjetiva e impracticable cuando el volumen a evaluar es muy alto. Ante el avance de industrias más tecnificadas, los sistemas de visión computacional y las herramientas de procesamiento de imágenes son una alternativa a la inspección tradicional. Este tipo de sistemas permite evaluaciones rápidas y objetivas, sirviendo a su vez como herramienta de registro y control de alimentos y procesos. En esta tesis, se propone emplear técnicas de procesamiento de imágenes para obtener información visual asociada a las características de calidad de productos agroindustriales. En particular, se estudiaron técnicas de procesamiento de imágenes en el estudio de la forma y color en dos alimentos: quesos tipo *Emmental* y frutillas.

En el caso de los quesos tipo *Emmental*, se estudió la formación de ojos durante el proceso de maduración, última y más larga etapa en la producción de este tipo de quesos. Los ojos son cavidades de gas dentro del queso que se forman durante la maduración como consecuencia de la fermentación de los azúcares presentes en la leche que forma el queso. A su vez, los ojos constituyen un aspecto de calidad del producto, los mismos deben tener un número, distribución y apariencia deseables. Por otra parte, desviaciones de este estándar de calidad pueden indicar problemas durante la producción de este alimento. En paralelo, se constata que durante la maduración los quesos sufren de cambios en su forma, alterándose su geometría. Como resultado de la investigación realizada en esta tesis, se pudo vincular los cambios en la geometría de las hormas con la formación de ojos en su interior. Para ello se realizó un seguimiento de

la maduración de hormas de queso, controlando sus dimensiones externas y su masa. En determinadas etapas de la maduración se cortó una horma para fotografiar su interior. Empleando técnicas de procesamiento de imágenes se obtuvo información del número y geometría de los ojos en una sección del queso. Aquí se utilizan diferentes técnicas de procesamiento de imágenes para la determinación del área de ojos en la sección. Se propusieron dos modelos que permiten estimar el volumen de ojos dentro de la horma: uno emplea la información de las imágenes y otro emplea las dimensiones geométricas de la horma junto a la densidad del queso y su masa. Se validaron ambos modelos al combinarlos y así estimar la masa del queso, contrastándola con la medida experimental. Los resultados indican que existe una relación entre el cambio de la altura máxima de la horma de queso y el volumen de ojos en su interior, lo cual implica que se puede conocer el estado de maduración del queso de forma indirecta y no destructiva.

El estudio realizado en frutillas implicó la determinación de zonas de oscurecimiento (“manchas pardas”) mediante el uso de fotografías. Las “manchas pardas” en frutillas maduras están asociadas a un producto envejecido y de menor calidad, por lo tanto, son un defecto del producto. En esta sección de la tesis, se propuso vincular la percepción visual del consumidor de frutillas frescas con sus correspondientes fotografías, con el fin de evaluar si es posible detectar las “manchas pardas” en las fotografías. Los resultados fueron que las “manchas pardas” son un aspecto detectable por consumidores de forma consistente tanto en el producto fresco como en las imágenes. A su vez, se evidenció que en el espacio de color HSV las “manchas pardas” pertenecen a un conjunto diferente que las zonas de la frutilla que no tienen este defecto. Por lo tanto, se puede emplear esta información en el diseño de un sistema de visión computacional para la detección de este tipo de defectos de forma automática, pudiendo emplearse en la categorización del fruto según su nivel de calidad o estado fisiológico.

A partir de los resultados obtenidos se concluye que las técnicas de procesamiento de imágenes son una herramienta que permite una evaluación rápida, objetiva y automática del aspecto de los alimentos. Empleando estas técnicas se puede obtener información que es de utilidad en el monitoreo de un proceso alimentario, tal como se realizó en los quesos. También, el procesamiento de

imágenes puede asistir en la evaluación del aspecto de los alimentos, como fue en el caso de las frutillas.

Palabras claves:

Procesamiento de imágenes, Monitoreo de procesos, Quesos formadores de ojos, Frutillas, Atributos de calidad.

ABSTRACT

Visual appearance of food products is one of the most relevant aspects for consumers when it comes to choosing a product. This is particularly relevant in the case of fresh produce such as fruits, vegetables, meat, dairy, among others. Consumers might believe a food product is inferior in quality when there are differences between their expectations and the actual aspect of the product. These differences might impact negatively on the commercialization of such products, even if they comply with other consumer expectations such as flavour, aroma, texture and so on. One method to categorize food according to its appearance is traditional inspection. However, this method is tedious, laborious, subjective and sometimes an impossible task to perform. New inspection technologies such as computer vision systems and image processing pose an alternative to traditional inspection. Fast and objective inspections can be achieved with these technologies, helping in process monitoring as well. In this thesis, image processing techniques were studied to obtain visual information related to agri-foods quality. In particular, shape and colour based techniques were studied on digital photographs of two food products: *Emmental* type cheese and strawberry.

Eyes formation in *Emmental* type cheese was studied during the last stage of cheese manufacturing, which is called ripening. Eyes are holes inside cheese that form during ripening as a by-product of fermentation of sugars in milk. Eyes are a quality aspect of cheese, their shape, number and appearance must comply with quality criteria. What is more, eye formation is sensitive to changes in cheese manufacturing, therefore eyes are an indicator of overall cheese quality. In addition to eye formation, cheese shape changes during ripening. In this thesis, a study on eye formation on *Emmental* type cheese during ripening was conducted. Experiments consisted on measuring cheese dimensions and weight during ripening. Also, cheeses were cut in half at different stages of the ripening process to register eye formation in photographs. Photographs were processed to obtain the number and area of eyes in cheese during ripening. Two different models were proposed to estimate eyes' volume in cheese: one model was based on image processing, and the other on cheese dimensions,

weight and density. Proposed models were validated by combining them to estimate cheese weight, and this estimation was contrasted with experimental measurements. Results showed there is a relationship between eye formation and geometrical changes in cheese during ripening. This means that eye volume inside cheese can be estimated using a fast and non-destructive method that involves using cheese volume and weight.

“Dark spots” in strawberries were studied using image processing on strawberries photographs. “Dark spots” are dark areas on strawberries’ surface, which appear on older strawberries and are a defect associated to a product of inferior quality. In this thesis, the relationship between consumer visual perception of “dark spots” on strawberries and photographs of such strawberries was studied. Experiments consisted on visual evaluation of strawberries by a panel, which was instructed to estimate the area of “dark spots” on strawberries. Then, the panel had to mark “dark spots” on strawberries photographs and estimate their extent. Both estimations were correlated to investigate if it is possible to detect “dark spots” on strawberries photographs. Results showed that consumers are consistent in “dark spots” detection on strawberries and in their photographs. Image processing of “dark spots” showed that they are statistically different in colour than non-affected areas. In addition, “dark spots” and non-spotted strawberry areas form two different clusters in HSV colour-space. Therefore, by photographing strawberries, image pixels can be classified as dark spots or not according to their intensity in HSV. This result may help in automatic “dark spot” detection and strawberry categorization according to its ripening and appearance.

Based on the results obtained in this thesis, image processing allows an automatic, rapid and objective evaluation of the appearance of food products. This information might help in process monitoring as in the case of cheese

manufacturing, and to evaluate strawberries' acceptability by consumers.

Keywords:

Image processing, Process monitoring, Eye-forming cheeses, Strawberries,
Quality attributes.

Lista de figuras

2.1	Espacio RGB	15
2.2	Espacios HSV/HSL	16
2.3	Espacio CIE L*a*b*	17
2.4	Escalado	19
2.5	Filtrado	20
2.6	Operadores morfologicos	22
2.7	Binarización	26
2.8	Detección de bordes: Algoritmos	29
2.9	Detección de contornos	31
2.10	Detección de círculos	33
2.11	Detección de regiones (Blobs)	34
3.1	Medidas geométricas del queso	41
3.2	Segmentación del queso respecto al fondo	45
3.3	Segmentación de ojos en imágenes “estándar”	46
3.4	Pretratamiento de imágenes “prueba”	48
3.5	Imágenes estándar	54
3.6	Evolución área de ojos y excentricidad	55
3.7	Resultados de segmentación de los algoritmos	56
3.8	Evaluación desempeño algoritmos	57
3.9	Área de ojos de algoritmos vs estándar	58
3.10	Evolución de la masa	61
3.11	Evolución de las medidas geométricas	62
3.12	Evolución volumen de ojos en queso	64
3.13	Volumen de ojos en queso en función de H	65
4.1	Imágenes de frutillas	73
4.2	Consistencia en estimación de área de manchas	80

4.3	Consistencia área de manchas en realidad e imágenes	80
4.4	Consistencia ubicación espacial de manchas	82
4.5	Muestra de consistencia ubicación espacial de manchas	83
4.6	Agrupamiento de regiones manchadas en espacios de color	84
4.7	Promedio regiones manchadas en espacios de color	85

Lista de siglas

A continuación se presentan las siglas empleadas en este texto junto a sus definicion

% HR Humedad relativa. Relación entre la presión parcial de vapor del agua respecto a la presión de vapor del agua en equilibrio en las condiciones de trabajo. **37, 40**

CIEXYZ Espacio de color estándar cuyas componentes XYZ son análogas a la respuestas de los conos LMS del ojo humano **14**

HSV Hue, Saturation, Value. Espacio de color cuyas componentes se corresponden a matiz, saturación y valor **16**

HSI Hue, Saturation, Lightness. Espacio de color cuyas componentes se corresponden a matiz, saturación y luminosidad **16**

MPx Mega píxeles **43**

RGB Red, Green, Blue. Espacio de color cuyas componentes son rojo, azul y verde **15**

ppp Píxeles por pulgada **43**

Tabla de contenidos

Lista de figuras	XIV
Lista de siglas	XVI
1 Introducción	1
1.1 Motivación y relevancia del tema	2
1.2 Objetivos, hipótesis y metodología	3
1.3 Estado del arte	5
1.4 Problemas abordados	7
1.5 Orientación al lector respecto al cuerpo del texto	9
2 Conceptos acerca del procesamiento de imágenes	10
2.1 Imagen digital	10
2.1.1 Imágenes fotográficas	11
2.2 Procesamiento digital de imágenes	12
2.2.1 Etapas del procesamiento de imágenes	13
2.3 Espacios de color	14
2.3.1 RGB	15
2.3.2 HSV/HSL	16
2.3.3 CIE $L^*a^*b^*/L^*u^*v$	17
2.4 Operaciones de pretratamiento	18
2.4.1 Cambio de escala	18
2.4.2 Filtrado	19
2.4.3 Operadores morfológicos	21
2.5 Segmentación por binarización	23
2.6 Bordes	26
2.7 Conectividad y etiquetado de regiones	29
2.8 Detección de contornos y curvas	30

2.8.1	Detección de contornos	30
2.8.2	Transformada de Hough y detección de curvas	31
2.8.3	Detección de blobs (regiones)	33
3	Detección de ojos en quesos tipo <i>Emmental</i>	35
3.1	Quesos: relevancia, características y estudios previos	36
3.1.1	Producción nacional	36
3.1.2	Características del queso tipo <i>Emmental</i>	37
3.1.3	Estudios previos del uso de imágenes en la detección de ojos en quesos	38
3.1.4	Propuesta de la tesis	39
3.2	Metodología	40
3.2.1	Obtención y conservación de los quesos	40
3.2.2	Masa, densidad del queso y geometría de las hormas	40
3.2.3	Estimación del volumen de ojos en el queso a partir de la geometría, masa y densidad	42
3.2.4	Obtención y procesamiento de las imágenes	43
3.2.5	Estimación del volumen de ojos de quesos a partir de imágenes	51
3.3	Resultados y discusión	53
3.3.1	Evolución de los ojos en los quesos y resultados de las imágenes “estándar”	53
3.3.2	Resultado de la detección semi-automática de ojos	55
3.3.3	Masa y medidas geométricas durante la maduración	60
3.3.4	Validación del las propuestas para la estimación del vo- lumen de ojos	62
3.3.5	Evolución del volumen de ojos durante la maduración y vinculación con medidas geométricas	64
4	Detección de “manchas pardas” en frutillas	67
4.1	Frutillas: relevancia, características y estudios previos	68
4.1.1	Frutillas a nivel nacional: producción y preferencia de consumo	68
4.1.2	Estudios previos del uso de imágenes en la evaluación de frutillas	69
4.1.3	Propuesta de la tesis	70

4.2	Metodología	71
4.2.1	Obtención y conservación de las frutillas	71
4.2.2	Entrenamiento del panel y evaluación visual	71
4.2.3	Obtención de las fotografías	72
4.2.4	Procesamiento de las imágenes	73
4.2.5	Procesamiento de datos	77
4.3	Resultados y discusión	78
4.3.1	Consistencia y capacidad del panel en la detección de “manchas pardas”	78
4.3.2	Evaluación del registro de “manchas pardas” en las fo- tografías	79
4.3.3	Análisis de diferencias en el color de zonas con y sin “manchas pardas”	83
5	Conclusiones y trabajo a futuro	86
5.1	Conclusiones generales	86
5.2	Conclusiones particulares	87
5.2.1	Estudio de quesos tipo <i>Emmental</i>	88
5.2.2	Estudio de “manchas pardas” en frutillas	88
5.3	Trabajo a futuro	89
	Referencias bibliográficas	91

Capítulo 1

Introducción

En la industria alimentaria es fundamental la producción de alimentos que cumplan los estándares de calidad ([Davies, 2000](#); [Gunasekaran, 1996](#)). La apariencia visual es un aspecto de calidad de relevancia en los productos alimenticios a partir del cual el consumidor puede decidir consumir o no el alimento ([Hutchings, 1977](#)). Una forma de evaluar el aspecto de los alimentos es mediante inspección tradicional, operación en la que personas calificadas evalúan los parámetros de calidad característicos de los productos alimenticios elaborados. Sin embargo, debido al costo, variabilidad y baja eficiencia asociada a la inspección tradicional, la industria se inclina hacia la automatización de esta tarea ([Zhang et al., 2014](#)). El uso de sistemas de visión computacional permite realizar tareas de inspección de forma rápida, no invasiva y objetiva, siendo de preferencia en líneas de producción y otras etapas del proceso donde es crítica esta tarea ([Brosnan and Sun, 2004](#)). Un sistema de visión computacional involucra adquirir una imagen o vídeo del producto o proceso a inspeccionar, procesar dicha imagen o vídeo y extraer información acerca de la escena inspeccionada ([Golnabi and Asadpour, 2007](#)).

Con la idea de obtener información del aspecto de productos alimenticios, en esta tesis se presenta el estudio de técnicas de procesamiento de imágenes en alimentos. Esta sección comienza presentando la motivación detrás de esta tesis y la relevancia del tema estudiado. Luego, se enuncian los objetivos e hipótesis que guiaron el trabajo realizado, junto a una breve descripción de la metodología empleada. Se prosigue con una revisión de la evolución del procesamiento de imágenes y su participación en la industria alimentaria. Posteriormente, se

presentan los productos agroalimentarios elegidos, sus características particulares y los problemas a resolver mediante el procesamiento de imágenes. Finalmente, la sección cierra con la descripción de la estructura del manuscrito.

1.1. Motivación y relevancia del tema

Esta tesis se enfoca en el estudio de técnicas de procesamiento de imágenes como método no destructivo para la determinación de aspectos de calidad de alimentos. El uso de estas técnicas surge como una nueva línea de investigación del grupo “Tecnologías Aplicadas a Procesos Alimentarios” del Instituto de Ingeniería Química de Facultad de Ingeniería (UdelaR). Este grupo se dedica al estudio de diversas técnicas para la mejora de los procesos de transformación y conservación de los alimentos. Algunas líneas de investigación del grupo consisten en el estudio de conservación de alimentos mediante procesos como envasado en atmósfera modificada y liofilización, revalorización de residuos agroindustriales mediante extracción asistida por ultrasonido, monitoreo no destructivo de procesos, y evaluación de propiedades reológicas en alimentos. La inclinación del grupo de “Tecnologías Aplicadas a Procesos Alimentarios” hacia el procesamiento de imágenes para obtener información de los alimentos surge como una expansión hacia el estudio de tecnologías de la industria 4.0 ¹ en las cadenas alimentarias, siendo esta tesis la primera del grupo sobre el uso de este tipo de tecnología en el monitoreo de alimentos. La motivación de esta tesis en particular es la implementación del procesamiento de imágenes como método de control de calidad de productos agroindustriales mediante la obtención de información visual externa que se vincule con el estado de calidad del producto.

El control y monitoreo de procesos alimentarios es de vital importancia para lograr procesos más eficientes y productos que cumplan los estándares de calidad (Sun, 2016; Patel et al., 2012). En muchos procesos, existen etapas en las que ocurren cambios visuales en los productos, asociándose ciertas características visuales a un determinado nivel de calidad. En este tipo de procesos,

¹El término industria 4.0 se refiere a la automatización y digitalización de los sistemas productivos tradicionales mediante la interconexión e intercambio de datos, la descentralización y la asistencia en la toma de decisiones por parte de sistemas no humanos.

el registro y procesamiento de imágenes puede ser de utilidad para controlar la calidad de los productos, registrar diferencias entre lotes y vincular cambios en el proceso productivo con alteraciones en los productos (Cubero et al., 2011). Mediante el procesamiento de imágenes se puede obtener información visual del proceso de una forma rápida y no invasiva (Brosnan and Sun, 2004). Esto permite su implementación en línea, obteniéndose resultados en tiempo real y logrando así un mayor control del proceso.

En la industria alimentaria, el uso de imágenes digitales y su procesamiento cumple funciones de control de calidad en productos terminados (Pathare et al., 2013), envasado (Ramirez et al., 2014), detección de defectos e impurezas (Du and Sun, 2004; Brosnan and Sun, 2004), control de trazabilidad de productos (Ribeiro et al., 2018), entre otras operaciones. El uso de sistemas que involucran la adquisición y procesamiento de imágenes permiten controlar operaciones en línea tales como el nivel de líquido en botellas, la presencia de objetos extraños, la colocación de etiquetas, la distribución de productos dentro de los envases, etc. Este tipo de control evita que productos defectuosos salgan al mercado, lo cual puede generar daños tanto a la salud de los consumidores como a la imagen de la empresa, además pérdidas debido a los costos asociados al retiro de estos productos del mercado. (Trienekens and Zuurbier, 2008). Respecto al uso de técnicas de procesamiento de imágenes en la agroindustria, las aplicaciones varían desde el monitoreo de plantaciones (Duggal et al., 2016) y control de ganado (Cangar et al., 2008), hasta el control de productos agroindustriales al momento de su ingreso a planta (Kondo, 2010) y su posterior procesamiento (Zareiforush et al., 2015).

1.2. Objetivos, hipótesis y metodología

El objetivo general de esta tesis es estudiar la aplicación de técnicas de procesamiento de imágenes a productos agroindustriales. Los objetivos específicos planteados al inicio de la investigación que dio origen a la tesis son los siguientes:

1. Estudiar las herramientas de procesamiento de imágenes digitales ópticas más adecuadas para la evaluación de aspectos visuales relacionados a la calidad de determinados productos agroindustriales.
2. Crear algoritmos de procesamiento de imágenes ópticas digitales usando las herramientas estudiadas en el objetivo (1) para la detección de la geometría y la evolución de los ojos que pueden registrarse en hormas de queso tipo Emmental cortadas en mitades.
3. Diseñar algoritmos de procesamiento de imágenes ópticas digitales usando los conocimientos adquiridos en el objetivo (1) con el fin de detectar manchas oscuras en frutillas maduras.

Las hipótesis en las que se basó esta investigación fueron:

1. La existencia de aspectos visuales asociados a la calidad del producto que pueden ser registrados en imágenes ópticas de dichos productos.
2. Los aspectos mencionados en la hipótesis (1) pueden ser registrados mediante implementos comerciales de uso residencial tales como cámaras fotográficas digitales y sistemas de iluminación LED y fluorescente.
3. Los aspectos mencionados en la hipótesis (1) y registrados según la hipótesis (2) pueden ser detectados mediante algoritmos de procesamiento basados en herramientas fácilmente implementables y existentes en librerías de libre acceso.

En todos los casos de estudio se registraron imágenes ópticas empleando una cámara digital comercial. En las imágenes se capturó una unidad de producto, con fondo contrastante e iluminación interior controlada. Se implementaron algoritmos de procesamiento digital de las imágenes obtenidas para cada producto en Python 3.7 ([Van Rossum and Drake, 2011](#)), empleando además las librerías Numpy 1.15.4 ([Harris et al., 2020](#)), OpenCV 3.4.1 ([Bradski and Kaehler, 2000](#)), SciPy 1.1.0 ([Virtanen et al., 2020](#)) y Scikit-image 0.14.1 ([Van der Walt et al., 2014](#)). Se evaluó el desempeño de los algoritmos de procesamiento de imágenes en la detección de los aspectos de calidad contrastando los resultados obtenidos con los indicados por técnicas de referencia. Dichas técnicas involucraron indicar de forma manual el aspecto de calidad estudiado sobre las

imágenes registradas y representan la respuesta del consumidor ante el producto. El grado de similitud entre los resultados del procesamiento de las imágenes y las técnicas de referencia determinó la eficiencia de los distintos algoritmos estudiados. Luego, se empleó la información obtenida con el procesamiento de imágenes para la estimación de los aspectos de calidad relevantes de los productos estudiados.

1.3. Estado del arte

La apariencia visual de los alimentos es el primer criterio empleado para su compra o consumo (Hutchings, 1977). Esto se debe a que los consumidores esperan cierto aspecto de los alimentos, siendo dicho aspecto asociado al nivel de calidad e inocuidad del alimento, junto a la experiencia de consumo. El uso de sistemas de visión computacional y procesamiento de imágenes permite controlar de forma no destructiva la calidad de los procesos alimentarios mediante inspección de los procesos y productos, evitando que productos defectuosos o que presenten un riesgo a la salud de los consumidores salgan al mercado (Pinder and Godfrey, 2012). A su vez, el uso de estas herramientas en los sistemas de inspección tiene múltiples ventajas respecto a la inspección tradicional: brinda resultados objetivos, tiene mayor capacidad de inspección y pueden emplearse de forma continua sin alteraciones en el criterio de decisión (Pinder and Godfrey, 2012; Davies, 2000).

La adquisición y procesamiento de imágenes digitales se remonta a inicios del año 1940 (Tönshoff et al., 1988), aunque la difusión de estas tecnologías en la industria alimentaria se inició con posterioridad. Las primeras aplicaciones se basaron en técnicas de imagenología mediante luz visible, infrarrojo y rayos X, con reportes acerca de la determinación de calidad en carne (Barten and Pfister, 1980; Knudsen, 1970), evaluación de frutas y vegetales (FINNEY JR, 1978), determinación de calidad de legumbres y cereales (Simpson, 1965). A partir de 1980, el uso del procesamiento de imágenes en la industria alimenticia fue en aumento gracias al desarrollo tecnológico sumado a la necesidad de procesos más estandarizados, niveles de producción más altos y productos con mayores niveles de calidad (Gunasekaran, 1996). A su vez, la inspección tradicional se volvió impracticable en las nuevas condiciones de producción, las

que requieren inspeccionar cientos de productos por segundo (Davies, 2000).

Al avanzar las tecnologías de imagenología, almacenamiento, transmisión y procesamiento de datos, la industria alimenticia acompañó esta tendencia, siendo una de las diez industrias que tienen mayor grado de incorporación de sistemas de visión artificial (Gunasekaran, 1996). En la actualidad, el uso conjunto de sistemas de visión y procesamiento de imágenes permite determinar propiedades externas e internas de los productos. Mediante imágenes en el rango de luz visible se suelen determinar propiedades externas tales como forma, tamaño, color, textura, brillo, distribución de elementos, entre otras características externas (Zhang et al., 2014). A su vez, mediante el conocimiento del proceso de transformación del alimento monitoreado, estas propiedades pueden vincularse con los cambios en las características de calidad del producto durante su procesamiento. Ejemplo de ello son la relación entre color y estado de madurez de hortalizas (Pathare et al., 2013), tamaño y deshidratación de alimentos (Sampson et al., 2014), textura y dureza en carne vacuna (Nunes et al., 2015; Li et al., 2001). Sistemas de imagenología más complejos permiten extraer información acerca del estado interno de los alimentos, tales como el uso de rayos X, imagenología hiperespectral, tomografía computada, resonancia magnética, etc. Existen múltiples ejemplos del uso de estas técnicas en el estudio de alimentos, sin embargo, a excepción del uso de rayos X, no se han difundido ampliamente en la industria como los sistemas de visión basados en luz visible (Chen et al., 2002). Además, si bien estas tecnologías pueden brindar información acerca de la estructura interna de los alimentos (la cual no puede obtenerse directamente mediante imágenes de luz visible), suelen ser mucho más costosas y complejas en su implementación que un sistema de visión que emplea imágenes ópticas.

Respecto al uso de algoritmos de procesamiento de imágenes, su evolución fue a la par de los desarrollos en software y hardware. En sus inicios estos algoritmos involucraron operaciones aritméticas sobre las imágenes, evolucionando a operaciones más complejas como la detección de bordes y formas (Tönshoff et al., 1988). El procesamiento de imágenes más simple se basa en operaciones a nivel de píxeles o grupos de píxeles, por lo que para lograr los resultados deseados se necesita conocimiento previo del sistema evaluado (Tillett, 1991). Al avanzar los sistemas de adquisición y procesamiento de imágenes, los algo-

ritmos de procesamiento de imágenes se volvieron más complejos y capaces de extraer mayor información de las imágenes. Estos avances fueron en paralelo con aumentos en la resolución y calidad de las imágenes adquiridas, evolucionando desde imágenes binarias, en escala de grises hasta imágenes a color y multiespectrales.

En la década de 1990 se presentan las primeras redes neuronales para la interpretación de imágenes, siendo el ejemplo más conocido la red LeNet ([LeCun et al., 1998](#)), la cual permite la identificación de caracteres. A partir de la década del 2010, gracias al aumento en la capacidad de cómputo y al gran volumen de datos disponibles, estas herramientas adquirieron mayor protagonismo en el área de visión artificial. Es así que se inició la era del “deep learning”, área del conocimiento que, mediante algoritmos de aprendizaje automático, busca la clasificación e interpretación de datos a partir de la evaluación de grandes volúmenes de datos del mismo tenor. En el caso del procesamiento de imágenes, el uso del “deep learning” permite comprender escenas, clasificar imágenes según su contenido, detección de objetos y extracción de información. Si bien estas estructuras tienen mejores resultados que el procesamiento de imágenes mediante técnicas más simples, los algoritmos de procesamiento de imágenes tradicionales siguen siendo empleados en tareas de inspección y control en la industria alimentaria, sobretudo en aquellas de baja variabilidad del aspecto a determinar y cuya escena es conocida.

1.4. Problemas abordados

Esta tesis se centró en la determinación de aspectos visuales de calidad de dos productos agroindustriales mediante el uso de técnicas de procesamiento de imágenes ópticas de nivel bajo y medio. La elección de productos agroindustriales se basa en el peso que tiene este sector en la economía uruguaya ([Uruguay XXI, 2020](#)) y la necesidad de soluciones técnicas de bajo costo para pequeños y medianos productores ([Castro-Fontoura, 2016](#)). Los productos a estudiar fueron quesos tipo *Emmental* y frutillas (*Fragaria x ananassa Duch.*), los cuales pueden ser beneficiados con una mejora en su selección ya que su aspecto visual es un factor relevante desde el punto de vista del consumidor

(Lewers et al., 2020; Wadhvani and McMahon, 2012).

En el caso de los quesos tipo *Emmental*, se estudió la estimación de la formación de ojos (cavidades de gas) mediante el vínculo entre el procesamiento de imágenes y la variación de su geometría externa. Los ojos constituyen un parámetro de la calidad de este tipo de quesos, y al mismo tiempo ofician de indicador de problemas en la producción de este alimento (Fröhlich-Wyder and Bachmann, 2007). Particularmente, la formación de ojos es un proceso que ocurre en la última y más larga etapa de producción de los quesos tipo *Emmental* y similares (Fox et al., 2017). Las actuales técnicas de estimación de formación de ojos en queso son destructivas y dependen de las habilidades del maestro quesero. Por otra parte, durante la formación de ojos ocurren cambios en la geometría del queso, los cuales son fácilmente detectables mediante un sistema de visión computacional. En esta tesis se estudiaron técnicas de segmentación para la detección automática de ojos en quesos tipo *Emmental*, estimando el área de dichos ojos en una sección de la horma de queso y vinculándola con la geometría de la horma durante la maduración del queso. El objetivo consistió en estimar de forma no destructiva la evolución de los ojos dentro del queso. Los resultados obtenidos en este trabajo muestran que existe un vínculo entre los cambios externos y la formación de ojos, ambos detectables y cuantificables mediante el procesamiento de imágenes.

El estudio de frutillas se centró en la determinación de la presencia de “manchas pardas”, zonas más oscuras en la superficie del fruto que se asocian a un envejecimiento o menor calidad del fruto (Lewers et al., 2020). Este defecto es indeseado en frutillas frescas, cuya apariencia es fundamental a la hora de la compra por parte de los consumidores. Si bien se reportan trabajos para la detección de otros defectos en frutillas tales como hongos (Liu et al., 2018) y daños en el fruto (Nagata et al., 2006), no se encontró en bibliografía una técnica para la detección y cuantificación de “manchas pardas” mediante el uso de imágenes. En este trabajo de tesis se propone evaluar el uso de fotografías para la detección de “manchas pardas” en frutillas maduras. Empleando como referencia la estimación de un panel sensorial, se buscó determinar si en las fotografías se detectan las “manchas pardas” y si se puede estimar su extensión y localización en los espacios de color. Los resultados obtenidos indicaron que existe una buena correlación entre la capacidad de detección de las “manchas

pardas” tanto en el producto fresco como en fotografías del mismo. A su vez, el análisis de las imágenes de las frutillas muestra que existen diferencias significativas en el color de las “manchas pardas” y que dichas diferencias pueden ser registradas en las fotografías. Estos resultados constituyen el primer paso en el diseño de un sistema de visión automático para la detección de este defecto en frutilla fresca.

1.5. Orientación al lector respecto al cuerpo del texto

Esta tesis se encuentra estructurada de la siguiente forma: portada, agradecimientos, tabla de contenidos, resumen, cuerpo de la tesis y anexos. El cuerpo de la tesis se divide en 5 capítulos:

1. Introducción
2. Conceptos acerca del procesamiento de imágenes
3. Detección de ojos en quesos tipo *Emmental*
4. Detección de “manchas pardas” en frutillas
5. Conclusiones y trabajo a futuro

Actualmente el lector se halla finalizando el primer capítulo, que consiste en la introducción y cuyo objetivo es presentar al lector una breve reseña del tema y trabajo de tesis. El segundo capítulo describe el marco teórico en el que se basa la tesis, abarca el concepto de imagen digital y espacios de color, las etapas del procesamiento típico de una imagen y las técnicas empleadas en esta tesis en el procesamiento de imágenes. Posteriormente, se presentan dos capítulos, cada uno correspondiéndose a cada producto estudiado en la tesis. A su vez, cada uno de estos dos capítulos se divide en tres secciones que se corresponden a una introducción al problema, metodología, resultados y discusión. La tesis finaliza con un último capítulo de conclusiones, que consiste en comentarios acerca del uso de técnicas de procesamiento de imágenes en los productos estudiados, aspectos a considerar en la implementación de estos sistemas, el aporte que pueden brindar estas técnicas a la industria alimentaria e ideas comprendidas dentro de trabajo futuro.

Capítulo 2

Conceptos acerca del procesamiento de imágenes

En este capítulo se presentan básicos relacionados al procesamiento digital de imágenes y las técnicas empleadas en esta tesis. Se define qué es una imagen digital, qué es el procesamiento de imágenes y los pasos generales en dicho procesamiento. Finalmente se describen dichos pasos junto a las técnicas más empleadas, su base teórica y sus propósitos. A menos que se indique una referencia, la descripción de las técnicas de procesamiento de imágenes se basó en la información brindada en [Gonzalez and Woods \(2002\)](#) y [McAndrew \(2015\)](#).

2.1. Imagen digital

Una imagen digital puede describirse como una función discreta y finita que recorre un campo bidimensional (x, y) , cuyos elementos se denominan píxeles. Un píxel es el elemento más pequeño que compone una imagen y adquiere un valor funcional discreto y finito. Una imagen digital puede obtenerse de dos formas: a partir de la interacción de una fuente de energía con el mundo físico, o a través de datos o imágenes preexistentes. La primera alternativa implica una fuente de energía que interactúa con el mundo físico (escena), un sensor que registra la respuesta de la escena ante la fuente de energía y un digitalizador que transforma la señal analógica del sensor en una señal digital ([Trussell and Vrhel, 2008](#)). Las fuentes de energía empleadas en la obtención de imágenes pueden ser ondas electromagnéticas u ondas mecánicas. La segunda alternativa

para obtener una imagen digital es mediante el procesamiento digital de datos u otras imágenes, generando así una nueva imagen que no es producto de la interacción de una fuente de energía con un medio físico.

Las imágenes digitales pueden clasificarse según sus píxeles en imágenes binarias, imágenes en escala de grises (monocromáticas) o imágenes a color. Estas categorías se diferencian en el rango de intensidades y el número de elementos que pueden poseer los píxeles, siendo esto un factor determinado en la obtención de la imagen. Una imagen binaria está compuesta por píxeles que poseen un valor escalar que puede tomar dos valores posibles: cero (0) o uno (1). Por convención, en una imagen binaria el valor 1 se corresponde con el píxel encendido o blanco, y el valor 0 con el píxel apagado o negro. Los píxeles de una imagen en escala de grises también tienen un valor escalar, pero el mismo puede tomar más de dos valores enteros de intensidad. El rango más usual está asociado a una codificación de 8 bits, con valores entre 0-255, asignándose el mínimo valor del rango al píxel apagado o negro y el máximo al píxel encendido o blanco. Las imágenes a color son aquellas en que los píxeles son elementos vectoriales, es decir que tienen más de un elemento y que cada elemento puede adquirir valores dentro de un rango de más de dos valores posibles. Las imágenes a color pueden considerarse una composición de n imágenes en escala monocromática, representándose como una matriz n -dimensional. Cada imagen monocromática (matriz bidimensional) se corresponde con la respuesta de la interacción de un sensor diferente con la fuente de energía. El ejemplo de imagen a color más conocido es el obtenido mediante la fotografía digital. Una imagen fotográfica digital de 24 bits se puede representar como una composición de tres imágenes monocromáticas codificadas en 8 bits, donde cada píxel es un vector de tres elementos cuyos valores se comprenden entre 0 a 255.

2.1.1. Imágenes fotográficas

En particular, una imagen fotográfica es la proyección de objetos del mundo real en un plano, obteniendo dicha proyección mediante el registro de la interacción de la radiación electromagnética visible con los objetos que se fotografían (Teubner and Brückner, 2019). En esta proyección se pueden superponer las imágenes de los objetos de la escena, perdiéndose la dimensión asociada a

la profundidad (Teubner and Brückner, 2019). Esta superposición implica una pérdida de la información relacionada a la profundidad, aunque esta información puede recuperarse mediante la superposición de imágenes de diferentes planos de la escena o mediante extracción de información de la profundidad que puede contener la imagen (Aharchi and Kbir, 2019). A diferencia de otros sistemas de imagenología, la fotografía tiene un estrecho vínculo con la visión humana. Debido a la naturaleza de las fotografías, todo el proceso de adquisición puede compararse con lo registrado por el ojo humano (Teubner and Brückner, 2019). Esto implica que este sistema de imagenología es el adecuado para el registro de imágenes como las que pueden visualizar los seres humanos.

2.2. Procesamiento digital de imágenes

Junto a la generación de imágenes digitales se desarrollaron herramientas que permiten su manipulación con dos objetivos principales:

- Mejorar su aspecto visual para su interpretación humana
- Adecuar su contenido para su procesamiento computacional

Entre los diversos autores no hay un consenso claro acerca de los límites del procesamiento digital de imágenes y otras áreas de procesamiento de información tales como la visión computacional, el *machine learning* y la inteligencia artificial. Sin embargo, hay cierto consenso acerca del grado de procesamiento de una imagen, el cual puede catalogarse según el tipo de resultado que se logra. Un procesamiento de bajo nivel involucra operaciones tales como reducción de ruido, mejora de contraste y acentuado de bordes, obteniéndose otra imagen como resultado. El procesamiento de nivel medio parte de una imagen a la cual se le aplican operaciones como la segmentación¹, descripción y clasificación de objetos, siendo el resultado una serie de atributos correspondientes a la imagen. Finalmente, un nivel alto de procesamiento implica comprender los atributos obtenidos mediante el procesamiento de nivel medio con el fin de emular la visión humana, es decir registrar una escena, diseccionarla y obtener información acerca de la misma con el fin de comprender cómo se compone.

¹En procesamiento de imágenes, segmentar implica particionar la imagen en agrupaciones de píxeles de acuerdo a un determinado criterio

2.2.1. Etapas del procesamiento de imágenes

El procesamiento de imágenes involucra una serie de pasos, los cuales dependen del objetivo a lograr y por ello no existe un único procedimiento. Sin embargo, un esquema de procesamiento típico involucra en primer lugar la adquisición de la imagen. Suele seguir una etapa de preprocesamiento, en la cual se realizan operaciones de nivel bajo para favorecer las etapas posteriores. Las operaciones en el preprocesamiento dependen de la imagen y lo que se desee obtener de ella, pero suelen involucrar operaciones que no alteran el su contenido de forma drástica. Ejemplo de dichas operaciones son una transformación afín, remoción de ruido o mejora del contraste.

Tras el preprocesamiento, la imagen obtenida puede someterse a procesos de mejora o de restauración, siendo el resultado de la mejora un aspecto subjetivo mientras que la restauración es una operación objetiva. También pueden aplicarse otras técnicas como procesamiento en el espacio de color o cambio en la resolución (nivel de compresión), esto con el fin de facilitar el acceso a la información que se puede hallar en un determinado espacio de color o en un cierto nivel de compresión. Tras esta serie de pasos, la salida es una imagen de la cual se puede extraer información si se procede con un procesamiento de nivel medio. En el procesamiento de nivel medio se realizan operaciones cuyo objetivo es obtener información acerca de la composición imagen. Algunas de estas operaciones pueden ser operaciones morfológicas para la extracción de las formas que representan la escena, segmentación para la disección de la imagen en sus componentes fundamentales, representación de los componentes segmentados y descripción de los mismos en categorías de información. El resultado de un procesamiento de nivel medio es información que se denomina atributos de la imagen.

Finalmente, en un nivel alto de procesamiento se toman los atributos de la imagen como entrada y el procesamiento involucra el etiquetado o categorización de estos atributos. Estos atributos pueden ser cuantitativos (por ejemplo, número de objetos segmentados) o cualitativos (por ejemplo, texturas o formas), empleando técnicas diferentes para cada uno de ellos. El objetivo de un nivel alto de procesamiento es el de reconocer los elementos que se hallaban en la imagen de partida y sobre la cual se aplicaron los pasos previos mencionados.

Las técnicas empleadas en este nivel de procesamiento involucran aprender a reconocer patrones y objetos a partir de los datos obtenidos de la imagen para así comprender la escena.

2.3. Espacios de color

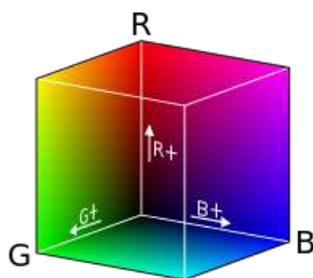
El color es una percepción humana de la interacción de la luz con los objetos ([Wyszecki and Stiles, 1982](#)). Los conos son células de la retina que permiten la visión a color, y presentan diferente sensibilidad a las distintas longitudes de onda que están asociadas a la percepción de los colores rojo, verde y azul. Debido a esto, los colores que percibimos son combinaciones de estos tres colores primarios aditivos, llamándose triestímulo a la cantidad de rojo, verde y azul necesaria para representar un determinado color. Los seres humanos distinguen los colores al comparar matiz, saturación y brillo del color. Se denomina matiz al color dominante, el cual está asociado a la longitud de onda mayoritaria de una combinación de ondas luminosas. La saturación se relaciona con la pureza del color o proporción de luz blanca en el color percibido, siendo un color saturado aquel que no se halla mezclado con luz blanca. El brillo es la percepción de la intensidad asociada a la falta de color (niveles de gris). Si se combinan el matiz y la saturación, se obtiene el concepto de cromaticidad.

La representación del color se realiza mediante espacios o modelos de color. Un espacio de color (generalmente tridimensional) consiste de un sistema de coordenadas y un subespacio en el que se representan los colores mediante puntos en dicho espacio. Los espacios de color suelen dividirse en dispositivo dependiente o dispositivo independiente. Los espacios dispositivo independiente son aquellos para los cuales existe una transformación no singular entre el espacio mencionado y el espacio de color CIEXYZ. El espacio CIEXYZ, descrito por la *Commission internationale de l'éclairage*, es el espacio que contiene los colores que una persona promedio es capaz de percibir. En caso de no existir una transformación no singular entre el espacio de color considerado y el CIEXYZ, el espacio es dispositivo dependiente. Los espacios dispositivos dependiente, no son capaces de reproducir todos los colores pertenecientes al espacio CIEXYZ, y suelen ser espacios diseñados para determinados dispositivos de registro o presentación de imágenes ([Trussell and Vrhel, 2008](#)). Existen múltiples espa-

cios de color, sin embargo, se presentarán tres de ellos: RGB, HSV/HSL y CIE $L^*a^*b^*$. Cabe mencionar que representar una imagen en un espacio de color diferente al cual fue registrada permite obtener una representación distinta de dicha imagen, sin alterar la información que la misma contiene. Es por ello que el cambio de espacio de color es una herramienta empleada en el procesamiento de imágenes, ya que representar la imagen en un espacio de color adecuado puede facilitar la extracción de la información deseada. En esta tesis se realizó esta aplicación reiteradas veces, siendo los resultados presentados en la figura 4.6 un ejemplo que permite visualizar los efectos del cambio de espacio de color sobre la información contenida en una imagen.

2.3.1. RGB

El espacio RGB es un espacio de color tridimensional basado en vectores asociados a los colores rojo, verde y azul (Red, Green, Blue). Es espacio de color basado en la adición de los colores primarios para la representación de los colores. Además, es un espacio orientado a hardware ya que sus canales están asociados a la representación de colores empleada comúnmente en pantallas. Por ello es el espacio empleado en la mayoría de los usos comerciales. No es un espacio perceptualmente uniforme, es decir, puntos de este espacio separados por una misma distancia pueden ser o no perceptualmente diferentes. Por lo tanto, no es un espacio adecuado para evaluar la percepción del color.



Los colores del espacio RGB se representan como vectores en un espacio cartesiano contenido en un cubo, como se ve en la figura 2.1. Los colores rojo, verde y azul son los vectores base del espacio, el negro está en el origen y el blanco en la esquina más lejana al origen.

Figura 2.1: Representación del espacio de color RGB. Imagen modificada de Wikipedia CC

2.3.2. HSV/HSL

Los espacios HSV (Hue, Saturation, Value) y HSL (Hue, Saturation, Lightness) son derivados del espacio RGB que fueron creados para que la representación de los colores fuera similar a la forma en que los humanos perciben el color. Al percibir el color, los humanos lo describen mediante sus atributos de matiz (hue), saturación y brillo. El matiz de un color es el atributo que describe el color puro que estamos percibiendo, por ejemplo: rojo, anaranjado, amarillo, azul. La saturación indica el nivel de mezcla de ese color puro con luz blanca, mientras que el brillo es un aspecto subjetivo asociado a la luminosidad del color. En los espacios HSV y HSL, la descripción del brillo se reemplaza con las componentes de valor y luminosidad, respectivamente.

Tanto el espacio HSV como el HSL son espacios cilíndricos (figura 2.2) y ambos describen al hue como el valor de la componente angular, siendo el ángulo 0° asignado al color rojo, 120° al verde y 240° al azul. A su vez describen la saturación como la componente radial, que tiene su máximo valor en el borde del cilindro, y el valor/luminosidad como el eje vertical.

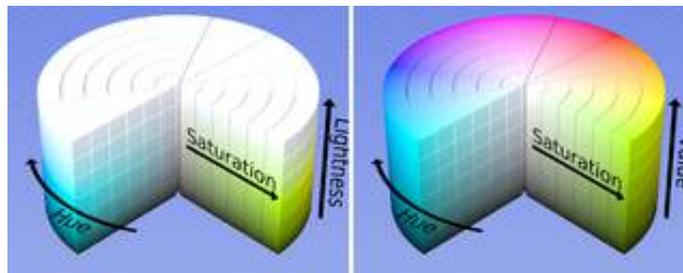


Figura 2.2: Espacios HSL (izquierda) y HSV (derecha) en su representación en coordenadas cilíndricas. Obtenida de Wikipedia CC

Sin embargo, existen ligeras diferencias entre el espacio HSV y HSL en cuanto a la interpretación de la saturación y brillo de un color. En el espacio HSV, la componente del valor es máxima para los colores saturados. En contraste, los colores saturados en el espacio HSL tienen una luminosidad correspondiente a la mitad del máximo valor posible de luminosidad. Es decir, en el espacio HSL, el nivel de saturación de un color no se ve afectado si el color se mezcla con blanco, pero se afecta al mezclarse con tonalidades de gris. En cambio, en el espacio HSV el nivel de saturación de un color se afecta tanto si se mezcla

con blanco o con cualquier tonalidad de gris.

2.3.3. CIE $L^*a^*b^*$ / $L^*u^*v^*$

El espacio de color CIE $L^*a^*b^*$ es un espacio dispositivo independiente que fue desarrollado para obtener un espacio de color que fuera perceptualmente uniforme. Esto quiere decir que si hay dos puntos perceptualmente diferentes separados una distancia d , cualquier otro par de puntos separados esa misma distancia d serán perceptualmente diferentes. Esto no ocurre en los espacios mencionados anteriormente ya que un determinado valor de distancia puede ser o no perceptualmente diferente dependiendo del punto considerado en dichos espacios. Su representación geométrica se muestra en la figura 2.3, en la que se aprecia la diferencia con los espacios de color antes presentados. En la figura 2.3 se ve que hacia el centro de la representación aumenta la luminosidad de los colores, mientras que en los bordes se representa lo que se corresponde a la matiz de los colores, expresados según su longitud de onda.

A su vez, el espacio CIE $L^*a^*b^*$ abarca todos los colores comprendidos en el CIEXYZ. Sus componentes son la luminosidad L^* y dos componentes de color, a^* y b^* . El componente a^* representa la cromaticidad a lo largo del eje verde-rojo, mientras que b^* representa la cromaticidad del eje azul-amarillo, teniendo valores negativos para los matices verde y azul, y positivos para el rojo y amarillo, respectivamente. Los valores en este espacio son relativos a un punto blanco de referencia, el cual debe ser informado.

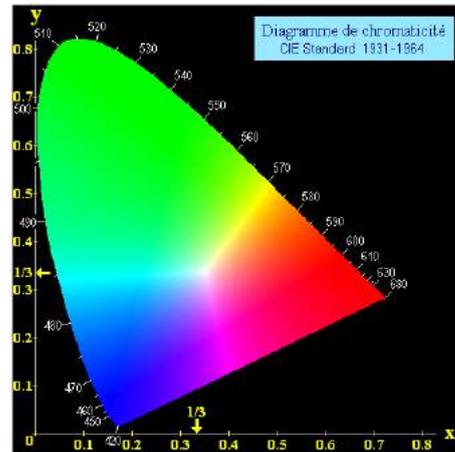


Figura 2.3: Espacio de color CIE $L^*a^*b^*$. Imagen obtenida de Wikipedia CC

2.4. Operaciones de pretratamiento

2.4.1. Cambio de escala

El cambio de tamaño o escalado es uno de los primeros pasos que se efectúan en el preprocesamiento. Escalar consiste en aplicar una transformación afín ² a la imagen original, obteniendo una nueva imagen del tamaño deseado y a cuyos píxeles se les asigna un valor de intensidad que depende de la imagen original. El método de asignación de los valores de los píxeles en la nueva imagen puede verse como la creación de una grilla con un número de elementos correspondiente al tamaño de la nueva imagen. Esta grilla se superpone ante la imagen original, por lo que sus elementos serán de tamaño mayor o menor a los elementos de la imagen original, según si se disminuye o incrementa el tamaño de la imagen respectivamente. Para determinar el valor de intensidad de los elementos de la grilla se realizan operaciones de interpolación en base a los valores de los píxeles de la imagen original.

Dos de los métodos de interpolación más usados son el método del vecino más cercano y la interpolación bilineal. El método del vecino más cercano consiste en asignar al nuevo elemento de la grilla el valor de intensidad del píxel de la imagen original que se encuentra a menor distancia del punto de la grilla evaluado. La interpolación bilateral considera los valores de intensidad de los cuatro vecinos más cercanos según su 4-conectividad, y el valor de intensidad resultante para la nueva imagen consiste en la interpolación bilineal de las intensidades de estos vecinos.

En la figura 2.4 se presentan los resultados de aumentar una imagen dos veces sus dimensiones empleando el interpolador de vecino más cercano (figura 2.4B) y bilineal (figura 2.4C). Observar que el uso del interpolador bilineal genera bordes más suaves que el uso del vecino más cercano. Se pueden proponer funciones más complejas como métodos de interpolación, lo cual puede brindar mejores resultados pero a costos computacionales más elevados. Una vez obtenidos los valores de los elementos de la grilla, estos se hacen corresponder con la matriz de la nueva imagen, lográndose el escalado.

²Una transformación afín es una operación matemática que en un espacio euclídeo preserva el paralelismo y la relación de distancias

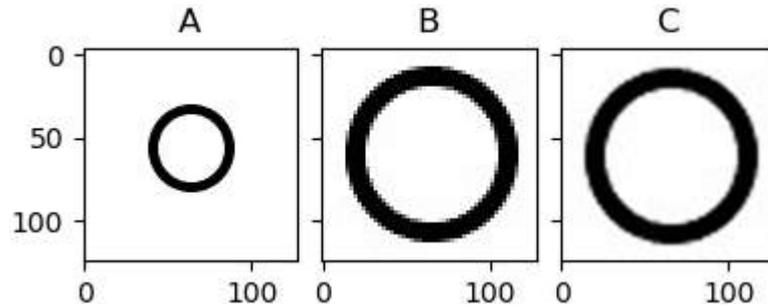


Figura 2.4: Ejemplo de aumento de tamaño. Imagen original (A). Aumento de escala 2 veces usando interpolación vecino más cercano (B) e interpolación bilineal (C).

2.4.2. Filtrado

El filtrado es una operación que elimina o resalta determinadas componentes en una señal y se suele emplear en el procesamiento de imágenes. Se denomina filtrado espacial a la aplicación de una función en cada píxel de la imagen original, obteniéndose una nueva imagen basada en la información que se preservó tras el filtrado. La particularidad del filtrado es que incluye la información de los píxeles vecinos que se hallan en un entorno determinado por el tamaño y forma del filtro (en general, cuadrados de lado impar). El filtrado espacial de una imagen consiste primero en escoger el tipo de filtro, es decir, la función a aplicar, el tamaño y la forma del entorno a considerar. Luego, el filtro se aplica sobre cada píxel de coordenadas x, y en la imagen original, y la respuesta será el valor del píxel en las coordenadas x, y de la imagen filtrada.

2.4.2.1. Filtros lineales

Los filtros se suelen dividir según su funcionalidad en lineales y no lineales. Un filtro lineal es aquel en que el valor resultante es una función lineal de todos los píxeles que son abarcados por el filtro. La operación que efectúan los filtros lineales es una convolución lineal, por lo que cumple con propiedades de linealidad, conmutatividad, asociatividad y separabilidad. Los filtros lineales más usados son el filtro de media (o promedio), el filtro de media ponderado, filtro gaussiano, filtro laplaciano, filtro de Sobel, entre otros. La ecuación 2.1 muestra el cálculo del valor de los píxeles $g(x, y)$ que resultan de filtrar una imagen f de tamaño $M \times N$ a la que se le aplica un filtro lineal de tamaño $m \times n$, con coeficientes w , siendo $a = (m - 1)/2$ y $b = (n - 1)/2$.

$$g(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x + s, y + t) \quad (2.1)$$

Ecuación 2.1: Expresión del resultado de aplicar un filtro lineal, siendo $g(x,y)$ el valor que adquiere el píxel x,y en la imagen filtrada.

Los filtros de media, media ponderado y gaussiano suelen emplearse en el suavizado de la imagen y reducción de ruido. El filtro de media opera efectuando el promedio de los valores de los píxeles acaparados por el filtro, mientras que el de media ponderada realiza un promedio ponderado. Un filtro gaussiano se basa en aplicar una función gaussiana normalizada 2-D de sigma predeterminado sobre los píxeles cubiertos por el filtro. Estos filtros suavizan la imagen y reducen el ruido al eliminar las componentes de alta frecuencia asociados a estos fenómenos. Este efecto se aprecia en la figura 2.5B, en la cual se aplicó un filtro gaussiano sobre la figura 2.5A.

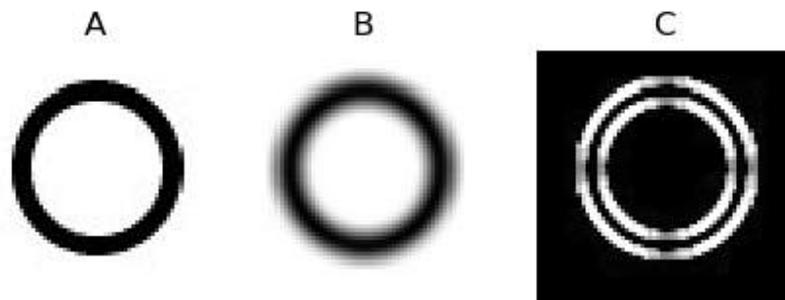


Figura 2.5: (A) Imagen original. (B) Resultado de aplicar un filtro suavizador (Gaussiano). (C) Resultado de aplicar un filtro de detección de bordes (Sobel).

En cambio, el filtro de Sobel y filtro laplaciano se emplean para resaltar bordes en la imagen mediante el uso de las derivadas de primer y segundo orden, respectivamente. Estos filtros preservan las componentes de alta frecuencia de una imagen, las cuales se asocian a bordes (zonas de cambios abruptos de intensidad). En la figura 2.5C se muestra un ejemplo de aplicar un filtro de Sobel a la figura 2.5A.

2.4.2.2. Filtros no lineales y morfológicos

Los filtros no lineales son aquellos en los que la función aplicada no es lineal. En particular, los filtros no lineales permiten conservar los bordes (componentes de alta frecuencia) mejor que los filtros lineales cuando se emplean en la reducción de ruido. Los ejemplos más conocidos son los filtros estadísticos de orden, bilateral, morfológicos, entre otros. Los filtros estadísticos de orden ordenan los valores de intensidad de los píxeles donde se aplica el filtro y escogen el valor de salida según un criterio determinado, por ejemplo la mediana, el máximo o el mínimo. Los filtros bilaterales dan como respuesta un promedio ponderado de los valores de intensidad de los píxeles del entorno. La función de ponderación de estos filtros puede depender de la intensidad y distancia de los píxeles del entorno.

Los filtros morfológicos se basan en operadores morfológicos, los cuales efectúan operaciones como la dilatación y la erosión sobre elementos de la imagen, empleando como base elementos estructurantes. En la siguiente subsección se profundiza en este tipo de operaciones y se describe su utilidad en el procesamiento de imágenes.

2.4.3. Operadores morfológicos

La matemática morfológica es una teoría propuesta por G. Matheron y J. Serra en 1964, y es una rama que estudia las estructuras geométricas basándose en la teoría de conjuntos. El empleo de herramientas de morfología matemática permite extraer los componentes de la imagen y así describir las formas y regiones que la componen ([Serra and Soille, 2012](#)). Las operaciones morfológicas en procesamiento de imágenes se basan en la interacción de dos conjuntos (imagen y elemento estructurante) mediante operaciones de la teoría de conjuntos como lo son la suma, resta, unión, intersección y complemento. Si bien en principio estas herramientas fueron desarrolladas para imágenes binarias, las mismas se extendieron a imágenes en escala de grises y a color.

Como se mencionó anteriormente, las operaciones de morfología matemática se basan en la interacción de un elemento estructurante y los conjuntos que forman la imagen. En imágenes binarias, se define al elemento estructurante

como una matriz binaria cuyo tamaño y distribución de elementos no nulos son parámetros a especificar (similar a un filtro espacial). Se toma como origen del elemento estructurante su centro, lo cual es relevante al aplicar las operaciones morfológicas. En cuanto a las operaciones que se pueden realizar con el elemento estructurante, la dilatación y la erosión son las operaciones básicas, a partir de las cuales se obtienen otras más complejas como apertura, clausura, transformación de localización, esqueletización, llenado de regiones, detección de bordes, entre otras.

La dilatación se define como la suma de todos los puntos con intersección no nula que se obtienen al desplazar la reflexión del elemento estructurante sobre los píxeles encendidos en el caso de una imagen binaria. El efecto de aplicar una dilatación es el de aumentar el tamaño del conjunto de píxeles encendidos según la interacción entre la imagen y el elemento estructurante. En la figura 2.6A se muestran los efectos de aplicar una dilatación sobre una imagen binaria (figura 2.6A) empleando un filtro morfológico rectangular. En la figura 2.6B se aprecia que el área ocupada por píxeles blancos aumentó, notándose en el aumento de tamaño de los puntos fuera del círculo y la desaparición de los puntos negros dentro del círculo (de tamaño menor al elemento estructurante).

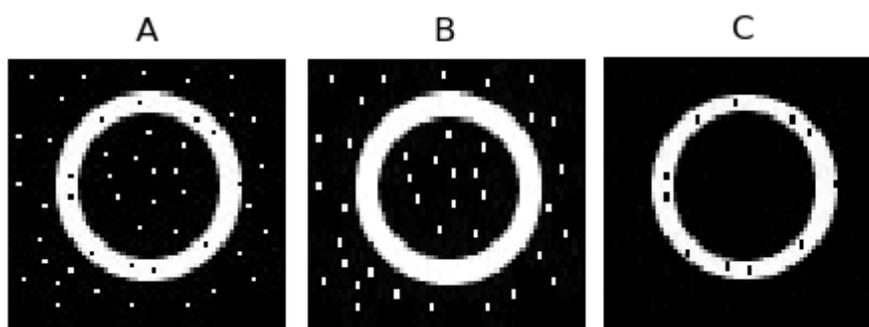


Figura 2.6: (A) Imagen binaria original. (B) Resultado de la dilatación. (C) Resultado de la erosión.

El resultado de la erosión son todos los puntos del elemento estructurante que están contenidos completamente en la imagen al desplazar el elemento estructurante sobre la imagen. El efecto de la erosión reduce el tamaño del conjunto imagen por el elemento estructurante. A diferencia de la dilatación, la erosión no es una operación conmutativa. Esto se debe a que esta operación elimina

conjuntos de la imagen de tamaño menor al elemento estructurante, por lo que una vez aplicada esta operación dichos conjuntos no pueden reconstruirse. En la figura 2.6C se muestran los efectos de aplicar una erosión sobre una imagen binaria (figura 2.6A), empleando un filtro con elemento morfológico de forma rectangular. Observar que en la figura 2.6C se eliminaron zonas blancas de tamaño menor que el filtro y se redujo el área ocupada por píxeles blancos. Esto último se aprecia claramente en un aumento del área de las zonas de píxeles negros que están rodeadas de píxeles blancos (figura 2.6C).

La apertura es el resultado de aplicar una erosión seguida de una dilatación sin cambiar el elemento estructurante. El uso de la apertura permite eliminar conjuntos más pequeños que el elemento estructurante, mientras que las estructuras más grandes son suavizadas al ser erosionadas y luego dilatadas. En cambio, la clausura es una dilatación seguida de una erosión, usando el mismo elemento estructurante. El efecto de la clausura es el de rellenar huecos que se hallan dentro de los conjuntos y que son más pequeños que los elementos estructurantes. Tanto la apertura como la clausura son operaciones idempotentes y duales, en el sentido de que realizar una apertura del conjunto del objeto equivale a realizar una clausura del fondo.

2.5. Segmentación por binarización

La binarización es una operación que permite segmentar píxeles al comparar el valor del píxel con un umbral y así clasificarlo. El ejemplo más simple de binarización es la binarización global, la cual implica escoger un único valor umbral para todos los píxeles de la imagen y clasificarlos como objeto o fondo dependiendo si el valor del píxel es mayor o menor a ese umbral, respectivamente. La elección de este umbral puede ser arbitraria o determinada por el histograma de la imagen a segmentar. El histograma de una imagen es un gráfico que muestra el número de píxeles de determinado nivel de intensidad en función de los niveles de intensidad.

2.5.0.1. Clasificación de algoritmos de binarización global

Según [Sezgin and Sankur \(2004\)](#), los algoritmos de binarización global se pueden clasificar según los siguientes métodos: basados en la forma del histograma, basados en agrupamientos, basados en entropía ³, basados en atributos de objetos y espaciales. En el siguiente párrafo se describen brevemente dichos métodos.

Los algoritmos basados en la forma del histograma buscan el valor de umbral en puntos de inflexión del histograma, o en puntos de mayor concavidad, o bien suavizando el histograma hasta hacerlo bimodal. Cuando el histograma es bimodal se emplean algoritmos basados en agrupamientos, y los métodos más populares son el método de Otsu ([Otsu, 1979](#)) y el ISODATA ([Ridler and Calvard, 1978](#)). Los algoritmos basados en entropía buscan el umbral que maximiza la entropía de la imagen binarizada o el umbral que minimiza la entropía cruzada entre la imagen original y el resultado de binarización, o bien un umbral que minimice la suma de la *fuzzy entropy*⁴. Los algoritmos basados en atributos escogen el valor umbral en el que se logran conservar atributos o medidas en la imagen binarizada comparándola con la original. Algunos de esos atributos pueden ser bordes, momentos, compacidad, textura⁵, entre otros. Finalmente, los algoritmos espaciales emplean información tanto de la distribución de los niveles de intensidad como del entorno de los píxeles, usando funciones de correlación, probabilidad de co-ocurrencia, dependencia lineal local, entre otras herramientas.

2.5.0.2. Binarización basada en el histograma de la imagen

En esta tesis las imágenes registradas consistieron en un objeto sobre un fondo contrastante e iluminación controlada. Es por ello que se lograron imágenes cuyo histograma era esencialmente bimodal. Es decir, los píxeles del objeto tendrán valores que se agruparán en una zona del histograma mientras que los

³La entropía de una imagen indica la cantidad de información que la misma contiene, es decir, el número de bits por píxel.

⁴En teoría de la información, la *fuzzy entropy* está asociada al cálculo de la entropía considerando la incertidumbre asociada a la falta de distinción entre los elementos que pertenecen y los que no pertenecen a un conjunto

⁵La textura de una imagen se refiere a la relación espacial entre los píxeles de una región de la imagen y está asociada a la percepción de patrones de textura visual.

píxeles del fondo tendrán valores que se agruparán en otra región del histograma. En esta situación, los métodos de binarización basados en el histograma de la imagen resultan adecuados para segmentar el objeto del fondo. Debido a que el histograma de las imágenes es bimodal, el valor umbral a escoger será aquel comprendido entre los dos modos del histograma. Como se mencionó anteriormente, dos de los algoritmos más empleados para determinar este valor de umbral son el algoritmo ISODATA y el algoritmo de Otsu.

El algoritmo ISODATA fue el primer algoritmo que propuso un método automático para obtener un valor umbral para la segmentación de una imagen. Este algoritmo busca el umbral de forma iterativa, suponiendo que el histograma de la imagen es bimodal. El algoritmo parte de un valor inicial de umbral, por ejemplo, el promedio de la imagen, y emplea dicho valor para separar el histograma de la imagen en dos clases. Luego, calcula el valor promedio de intensidad de las dos clases obtenidas, resultando en un nuevo valor umbral. El algoritmo continúa iterativamente hasta lograr la convergencia del valor umbral.

La idea que propone el algoritmo ISODATA puede plantearse como un problema de optimización mediante una función objetivo que logre un valor de umbral óptimo. El algoritmo de Otsu es un método que halla este umbral óptimo de forma automática, basándose en la varianza entre clases. En su versión simple, este algoritmo supone que el histograma de la imagen es bimodal. Según este método, el cálculo del umbral se basa en maximizar la varianza entre clases del histograma, las cuales en su caso simple son dos: objeto y fondo. La varianza entre clases se define como la suma ponderada de las varianzas de cada clase multiplicadas por su probabilidad. En un histograma bimodal, cada clase se corresponde a los píxeles asociados al objeto y al fondo, mientras que la probabilidad de cada clase es número de píxeles por clase divididos el número total de píxeles de la imagen.

La figura 2.7C muestra el resultado de aplicar el algoritmo de Otsu para la binarización de una imagen en escala de grises (figura 2.7A). En el ejemplo, si bien el resultado es bastante bueno, debido a la forma del histograma (figura 2.7B) no se logra segmentar correctamente todos los objetos. Esto se debe a que el histograma se extiende en un rango de intensidades acotado y esto re-

sulta en un bajo contraste entre el objeto más claro y el fondo blanco. Una solución consiste en corregir el histograma de la imagen para que abarque un rango mayor de niveles de intensidad. Para ello, se genera una nueva imagen con un rango mayor de intensidades al asignar a cada pixel un nuevo valor en función del histograma de la imagen original. La figura 2.7D muestra el resultado de hacer una corrección del histograma de la imagen previo a aplicar el algoritmo de Otsu, logrando una mejor segmentación respecto a la figura 2.7C.

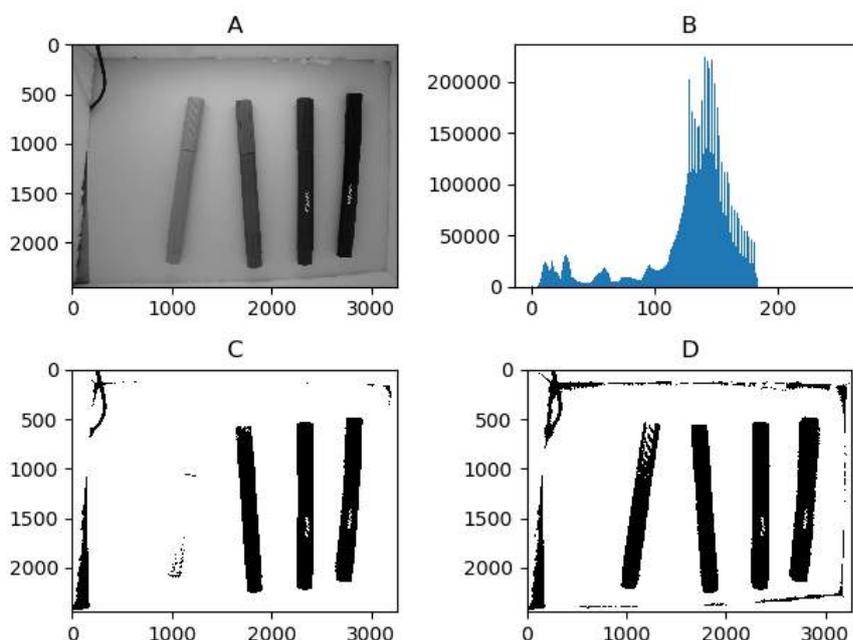


Figura 2.7: (A) Imagen original. (B) Histograma de (A). (C) Binarización de (A) empleando el método de Otsu. (D) Binarización de (A) empleando una mejora de histograma (CLAHE)⁶seguida del método de Otsu.

2.6. Bordes

Cambios de color o intensidad en las imágenes son elementos de importancia en la percepción e interpretación de la imagen. Elementos como bordes o

⁶CLAHE: Contrast Limited Adaptive Histogram Enhancement. Método de mejora del histograma que emplea información regional de la imagen para el cálculo de los nuevos valores de los píxeles y que limita el contraste de las imágenes resultantes para evitar problemas de sobresaturación.

contornos son parte de la descripción primaria de la imagen, la cual es una descripción que contiene mucha información de dicha imagen (Marr and Poggio, 1979). Se puede definir como borde en una imagen al conjunto de píxeles conectados según un determinado criterio de conectividad, para los cuales a nivel local la primera derivada bidimensional en dicho píxel es mayor a un umbral predefinido. Otra definición equivalente de borde implica a aquellos píxeles conectados cuya segunda derivada es nula en dichos píxeles.

Los métodos más populares para la detección de bordes consisten en la aplicación de operadores (filtros) sobre la imagen, los cuales se basan en el cálculo de la primer o segunda derivada a nivel local de la imagen (Savant, 2014). Este procedimiento resulta en una imagen donde los bordes y elementos de alta frecuencia son destacados. Posteriormente, se suele realizar una binarización de la imagen para continuar con la detección de contornos. Algunos de los operadores más populares son el operador de Prewitt, Sobel y el laplaciano de una gaussiana, los cuales se describen a continuación.

Los operadores de Sobel y Prewitt se basan en el cálculo de los gradientes de la imagen, pudiéndose aplicar para detectar bordes horizontales, verticales y diagonales. Ambos operadores están diseñados de forma que aplican un filtro que suaviza la imagen para disminuir el efecto del ruido en la detección de bordes. La diferencia entre ambos reside en que el operador de Sobel asigna un peso mayor a los componentes del filtro que suavizan la dirección central donde se aplica el filtro.

Otro operador ampliamente usado en la detección de bordes es el laplaciano de una gaussiana, que es la combinación de un filtro gaussiano con un operador laplaciano. Esta combinación logra suavizar la imagen mediante un filtro gaussiano de forma de eliminar ruido, mientras que el operador laplaciano detecta los bordes mediante el cálculo de los puntos donde se anula derivada segunda de la imagen. El uso del laplaciano tiene la ventaja de ser isotrópico a diferencia de los operadores basados en el gradiente. Otro aspecto interesante de este operador es que ciertos aspectos de la visión humana pueden modelarse con la definición matemática del mismo (Marr and Ullman, 1981). Este método es empleado en el algoritmo de Marr-Hildreth, el cual propone tres pasos para la detección de bordes:

1. Aplicar un filtro gaussiano
2. Convolucionar el resultado con un operador laplaciano
3. Detectar los puntos de *zero-crossing*⁷.

El inconveniente de aplicar un operador de forma directa sobre la imagen es la detección de bordes falsos, los cuales son causados por la presencia de ruido en la imagen o debido a la diferencia de escalas entre el operador y los bordes en la imagen. El detector de bordes de Canny busca solucionar estos inconvenientes, ya que si bien la detección de bordes se basa en el uso de filtros, se implementa de forma tal de solucionar los problemas causados por ruido y la diferencia de escalas. El detector de bordes de Canny permite detectar bordes en una imagen mediante la aplicación de un set de filtros de diversas orientaciones y trabajando en distintas resoluciones de la imagen, combinando el resultado de estas operaciones en la imagen final (Canny, 1986). Este procedimiento logra detectar bordes en diferentes orientaciones y a distintas escalas, disminuyendo la detección de puntos de falsos bordes y dando una buena ubicación de los puntos de borde. En el resultado final muestra solo una marca por borde y no todas las que pudo haber detectado al cambiar los tamaños de filtros y resolución. El filtro que emplea este método se basa en el uso conjunto de gradientes (para detectar los bordes) y los puntos de *zero-crossing* para obtener una mejor ubicación de los mismos.

En la figura 2.8 se ejemplifican tres detectores de bordes: 2.8B es el resultado de aplicar un filtro de Sobel bidireccional sobre la figura 2.8A, la figura 2.8C es el resultado de un filtro laplaciano y 2.8D es el resultado del detector de bordes de Canny. Si bien el detector de bordes de Canny logra resultados muy buenos en comparación al uso de filtros para la detección de bordes, el desempeño del algoritmo depende de la correcta elección de los parámetros que el algoritmo requiere. En este sentido, el uso de los detectores de bordes basados en filtros suele ser más sencillo.

⁷Punto de zero-crossing son los píxeles que cumplen una de las siguientes afirmaciones: (a) Tienen valor negativo y uno de sus vecinos de 4-conectividad tiene valor positivo; o (b) Tienen valor nulo y tienen un vecino de 4-conectividad con valor positivo y otro con valor negativo.

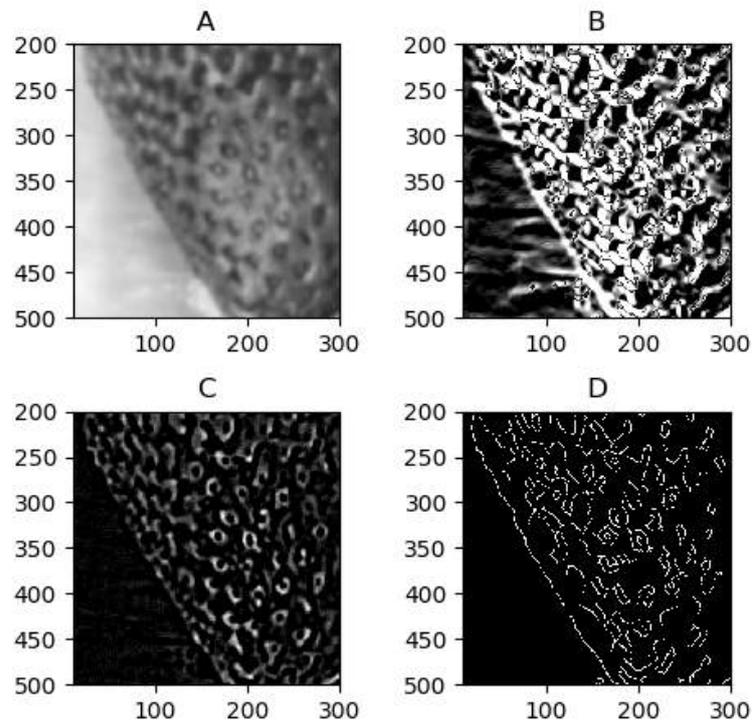


Figura 2.8: (A) Imagen en escala de grises. Resultados sobre (A) de: binarización de un filtro bidireccional de Sobel (B), filtro laplaciano (C), y detector de bodes de Canny (D).

2.7. Conectividad y etiquetado de regiones

El etiquetado de regiones permite agrupar píxeles con determinada conectividad⁸ como pertenecientes a un mismo grupo al asignarles un valor común. De esta forma, se generan regiones que se asocian a los objetos que componen la imagen. Mediante el etiquetado se puede contabilizar, localizar y determinar las propiedades geométricas de las regiones etiquetadas. Así, se obtiene información de la composición de la imagen, permitiendo su posterior interpretación.

Existen diversos algoritmos para realizar el etiquetado de regiones (He et al., 2017). Uno de los más sencillos consiste en el etiquetado secuencial o etiquetado por resolución de equivalencias, el cual es un algoritmo en el que se recorre la imagen y se asignan etiquetas a los píxeles no nulos. Al hallar un píxel con

⁸En imágenes, la conectividad de los píxeles indica la ubicación de un píxel respecto a sus píxeles vecinos

valor considerado como objeto (1 en el caso de imágenes binarias), se evalúa su conectividad, y en función de ello se le asigna una nueva etiqueta o la etiqueta de sus vecinos. En este algoritmo se evalúa la conectividad de cada píxel respecto a sus vecinos superiores y hacia la izquierda, según el tipo de conectividad elegida (4 u 8 conectividad). Por ello, puede ocurrir que un píxel tenga dos o más vecinos con diferentes etiquetas. Esto se resuelve registrando estos casos como equivalentes y en un siguiente barrido de la imagen se reemplazan las etiquetas equivalentes con un nuevo valor común. Otro algoritmo que permite el etiquetado de regiones se basa en el llenado de regiones. En este caso, se buscan píxeles no nulos sin etiquetar y a partir de ellos, se etiquetan todos los píxeles conectados bajo la misma etiqueta.

2.8. Detección de contornos y curvas

Tras la segmentación o detección de bordes, se suele obtener una imagen binaria como salida, la cual es una simplificación de la imagen original. Sin embargo, no se tiene información acerca de la relación entre dichos píxeles, por ejemplo, si describen algún tipo de elemento geométrico, su ubicación u orientación en la imagen. Este tipo de información permite tener una descripción primaria de la imagen analizada, por lo que resulta de interés conocerla. A su vez, existen diversos algoritmos y métodos para vincular los píxeles segmentados entre sí y obtener descriptores que permitan identificar contornos y formas. Debido a esto, se describen a continuación los algoritmos empleados en esta tesis para la detección de contornos y de formas geométricas sencillas. En particular, se emplearon las implementaciones disponibles en OpenCV para la detección de contornos, detección de círculos y detección de blobs (regiones).

2.8.1. Detección de contornos

Una manera de detectar contornos se basa en obtener una secuencia de coordenadas que describe la ubicación de un grupo de píxeles encendidos conectados mediante el algoritmo de seguimiento de bordes. El algoritmo empleado en esta tesis es el disponible en OpenCV, el cual se denomina `findcontours` y está implementado según el trabajo de [Suzuki et al. \(1985\)](#). Esta implemen-

tación extiende la idea del seguimiento de bordes, algoritmo que trabaja con los píxeles encendidos (blancos) conectados de una imagen binaria, y que diferencia bordes externos e internos. La implementación de [Suzuki et al. \(1985\)](#) además identifica cada borde de forma única y añade un procedimiento para conocer a que borde pertenece cada píxel. El resultado del `findcontours` da un vector de puntos para cada contorno identificado, y los contornos identificados se ordenan de forma jerárquica, con los menores índices asignados a los contornos más externos.

La figura 2.9C muestra un ejemplo de detección de contorno tras la binarización de una imagen a color de una placa de Petri. A partir de estos contornos se podría realizar el recuento de las colonias en la placa de Petri, conocer el tamaño de las mismas y color de las mismas, entre otros parámetros de interés.

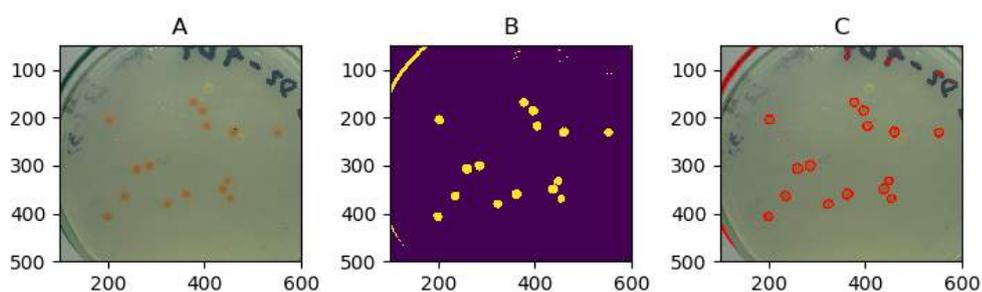


Figura 2.9: (A) Imagen original. (B) Resultado de binarización por umbralización de (A). (C) Detección de contornos de (B) y representación sobre (A).

2.8.2. Transformada de Hough y detección de curvas

A diferencia de la detección de contornos descrita en la sección anterior, el uso de la transformada de Hough para la detección de formas se basa en un enfoque de búsqueda global de estructuras. Este método, propuesto por Paul Hough en 1962, se basa en localizar puntos que componen figuras que pueden ser parametrizadas, tales como líneas, círculos, elipses, etc. La idea detrás de este método es transformar el espacio cartesiano en otro espacio basado en formas parametrizadas de las figuras a detectar.

La forma más conocida de la transformada de Hough es su aplicación a la detección de líneas. En este enfoque, se transforma el espacio de coordenadas

cartesianas en coordenadas polares, de forma que una línea puede representarse como $x\cos(\theta) + y\sin(\theta) = \rho$ (Ballard, 1981). El algoritmo se basa en dividir el espacio en conjuntos de (ρ, θ) . Luego, el algoritmo busca puntos en la imagen dentro de este espacio y se contabiliza el número de puntos que pertenecen a cada conjunto de (ρ, θ) en una matriz de acumulación de dimensiones $[\rho \times \theta]$. Tras recorrer la imagen, cada coordenada de la matriz de acumulación indica el número de puntos que están asociados a un par (ρ, θ) . La detección de rectas se basa en obtener los pares (ρ, θ) que tienen una cantidad de puntos superior a un valor predeterminado. La idea es hallar las líneas más relevantes en una imagen mediante la detección de un número significativo de puntos alineados que comparten (ρ, θ) , y cuya detección es más sencilla mediante la inspección del espacio de la imagen en coordenadas polares.

Debido a su versatilidad, se puede emplear la transformada de Hough para la detección de puntos que pertenezcan a una función $g(v, c) = 0$, donde v es un vector de coordenadas y c un vector de coeficientes. De esta forma, se puede adaptar la transformada de Hough para la detección de círculos. En este caso, los puntos de un círculo cumplen $(x - C_1)^2 + (y - C_2)^2 = C_3^2$, siendo C_1, C_2, C_3 los parámetros que formarán un subespacio 3D en el cual se contabilizarán los puntos pertenecientes a círculos con centro C_1, C_2 y radio C_3 . Al igual que para la detección de líneas, se fija un umbral a partir del cual se considera que un conjunto C_1, C_2, C_3 posee suficientes puntos de forma que describe un círculo. De esta forma es que se obtiene la variante conocida como transformada de Hough circular. Con este mismo enfoque puede emplearse esta transformada para detectar elipses al usar la expresión de la función que describen los puntos pertenecientes a una elipse.

La figura 2.10C muestra el resultado de la detección de círculos usando la transformada de Hough circular sobre el canal de saturación (figura 2.10B) de una imagen a color. Observar que en figura 2.10C hay círculos no detectados junto a numerosos casos de falsos positivos. Esto resulta de la elección de los parámetros de búsqueda del algoritmo, presencia de ruido en la imagen, entre otros factores que afectan el resultado del algoritmo.

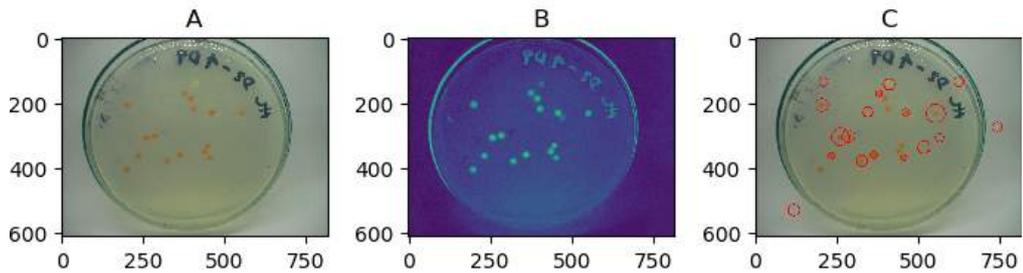


Figura 2.10: (A) Imagen original. (B) Canal de saturación de (A). (C) Detección de círculos de (B) y su representación sobre (A).

2.8.3. Detección de blobs (regiones)

Un blob es una región de una imagen en la cual algunas propiedades son prácticamente constantes. La detección de este tipo de regiones permite separar píxeles por propiedades tales como color, área, compacidad, convexidad, etc. El algoritmo empleado (`simpleBlobDetector`) se centra en la detección de regiones con un determinado nivel de circularidad. Este algoritmo tiene la particularidad de ser bastante robusto y considerar varios aspectos a la hora de seleccionar los blobs, filtrando los resultados por color, área, circularidad, relación de inercia y convexidad.

`simpleBlobDetector` efectúa una serie de pasos sobre la imagen de entrada para detectar las regiones que cumplen con los parámetros ingresados. Primero, binariza la imagen aplicando una serie de umbrales entre un valor mínimo y máximo determinados. Luego, extrae los componentes conectados usando la función `findContours` y calcula sus centroides. Por último, el algoritmo agrupa los centros de las binarizaciones obtenidas en el primer paso según un criterio de distancia mínima, para así asignarlos a un blob, devolviendo como salida las coordenadas de los blobs y su tamaño.

La figura 2.11 es un ejemplo de aplicación del detector de blobs, cuyos parámetros se fijaron para la detección de círculos brillantes del canal de saturación. Al comparar esta figura con los resultados obtenidos para la transformada de Hough sobre la misma imagen, se aprecia que el algoritmo `simpleBlobDetector` fue más preciso en la detección de las colonias sobre la placa.

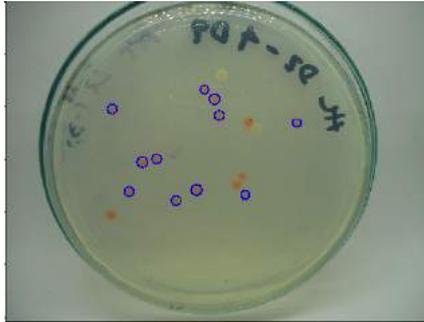


Figura 2.11: Resultado de la detección de blobs sobre la imagen 2.10B, representada sobre 2.10A.

Capítulo 3

DetECCIÓN DE OJOS EN QUESOS TIPO *Emmental*

Este capítulo se centra en la estimación de la evolución del volumen de ojos en quesos tipo *Emmental* mediante una propuesta no destructiva basada en medidas geométricas y la masa de las hormas. Esta propuesta se validó empleando un modelo que permite estimar el volumen de ojos a través del procesamiento de imágenes de los quesos cortados longitudinalmente. Para la obtención de los datos necesarios, se realizó un ensayo de maduración con hormas de queso tipo *Emmental* y durante dicho ensayo se registraron las dimensiones externas y masa de las hormas. En paralelo, cada 5 días se cortaba una horma por la mitad para fotografiar su interior. Empleando las técnicas de procesamiento de imágenes presentadas en el capítulo 2, se extrajo información geométrica de los ojos en las imágenes. Con el área y ubicación de los ojos en las imágenes, se estimó el volumen de ojos en el queso y se empleó esta información para validar la propuesta de estimación no destructiva del volumen de ojos en quesos.

Este capítulo se divide en tres grandes bloques, en los cuales se presenta el problema a estudiar, la metodología empleada y los resultados obtenidos. En el primer bloque se presenta la relevancia de los quesos en Uruguay, la elaboración de quesos tipo *Emmental* junto a la importancia de los ojos, una revisión bibliográfica de la determinación de ojos en quesos y la propuesta de esta tesis. En el segundo bloque se describe el registro de las medidas externas de las hormas de queso, la obtención y análisis de las imágenes del interior de los quesos y los modelos propuestos para vincular ambos datos. El último

bloque consiste en la presentación de los resultados obtenidos y su discusión en función de los objetivos planteados.

3.1. Quesos: relevancia, características y estudios previos

El queso es un producto alimenticio que se obtiene mediante coagulación de la leche. Existen múltiples variedades de quesos que se diferencian según el método de coagulación, el tipo de leche empleada, el proceso de elaboración, las características reológicas del producto, entre otros factores. En Uruguay los quesos son productos de alto consumo debido a factores culturales, su calidad y disponibilidad. Al ser un producto de alto valor económico, tecnologías que permitan su mejora son valoradas por la industria. En esta sección se presentará la relevancia de este alimento agroindustrial, sus características, una breve revisión bibliográfica acerca del procesamiento de imágenes para la detección de ojos en quesos y la propuesta de esta tesis.

3.1.1. Producción nacional

A nivel nacional, el queso es un producto de alto consumo, estimándose para el año 2018 en 8 kg/hab/año según datos obtenidos del Ministerio de Ganadería, Agricultura y Pesca ([Oficina Estadísticas Agropecuarias, 2019](#)). La producción de quesos se realiza principalmente en los departamentos de San José, Colonia, Canelones y Florida. El consumo interno representa 49,4 % del total producido, teniendo una tendencia al alza en los últimos años ([Oficina Estadísticas Agropecuarias, 2019](#)). En 2015, la producción de queso destinada al mercado interno fue de 21.603 ton mientras que en 2018 fue de 28.366 ton, destacando la relevancia de este producto a nivel nacional. En esta tesis se trabajó con queso tipo *Emmental*, un queso semiduro de pasta cocida que se obtiene a partir de leche bovina, y que pertenece al grupo de “quesos formadores de ojos” junto al Edam, Maasdam, Colonia, entre otros.

3.1.2. Características del queso tipo *Emmental*

La fabricación de queso tipo *Emmental* involucra la coagulación de la leche, corte y cocción del coágulo, prensado en hormas cilíndricas, salado y finalmente la maduración (Fox et al., 2017; Law and Tamime, 2011). Durante la maduración, este queso permanece en cámaras de temperatura y humedad controlada durante un periodo de seis semanas a doce meses, dependiendo del tamaño de la horma, la temperatura de la cámara y las características sensoriales que se desean en el producto final (Fox et al., 2017; Law and Tamime, 2011). En particular, la maduración de quesos tipo *Emmental* se realiza en dos etapas. La primera etapa se denomina maduración en cámara fría y se realiza a 10–15 °C y 90 % HR (humedad relativa) durante 10 a 14 días. Le sigue la maduración en cámara caliente, la cual se realiza a 16–24 °C y 80 % HR durante un mínimo de cuatro semanas, pudiéndose extender por varios meses (Fox et al., 2017).

En la etapa de maduración de quesos ocurren procesos bioquímicos que dan las características de sabor, textura, aroma y color característicos de la variedad (Fox et al., 2017; Law and Tamime, 2011). El queso tipo *Emmental* se caracteriza por tener burbujas de gas en su interior, las cuales se conocen como “ojos”. Los ojos se forman durante la maduración en cámara caliente y son resultado de la fermentación microbiana de la lactosa y otros azúcares disponibles (Fox et al., 2017). A su vez, los ojos son una característica de calidad importante que influencia la elección del consumidor. Se desea que los ojos sean redondeados, brillantes, tengan una distribución y tamaño determinados (US Department Agriculture, 2001). Por otra parte, la evaluación de los ojos permite diferenciar tanto los distintos niveles de calidad en el queso como la detección de problemas en la producción tales como contaminación o cambios en las condiciones de maduración (Fröhlich-Wyder and Bachmann, 2007). A su vez, se observa que durante la maduración de este tipo de quesos las hormas se deforman por dos efectos contrapuestos. Por un lado, el queso fluye ante la fuerza peso debido a su comportamiento viscoelástico; por otro lado, la formación de ojos provoca la expansión de la matriz queso debido a la presión que ejercen las burbujas de gas en el interior de la horma.

3.1.3. Estudios previos del uso de imágenes en la detección de ojos en quesos

En la literatura se han reportado con anterioridad el uso de diversos métodos de adquisición y procesamiento de imágenes con el fin de monitorear la formación de ojos en quesos, estimar su número, tamaño y geometría. Diversos investigadores han reportado el uso de técnicas de imagenología como la tomografía computada (Schuetz et al., 2013), rayos-X (Kraggerud et al., 2009), resonancia magnética (Huc et al., 2014b) y ultrasonido (Eskelinen et al., 2007) para la detección no destructiva de cavidades de gas dentro de las hormas de quesos. Si bien se han reportado buenos resultados con el uso de estas técnicas, su empleo comercial no es atractivo debido a su complejidad, limitaciones y/o alto costo.

A su vez, el uso de imágenes ópticas se ha reportado como una herramienta de control de calidad visual externa en diversas variedades de queso (Lukinac et al., 2018; Caccamo et al., 2004; Wrangborg and Ardö, 2007). La propuesta de Caccamo et al. (2004) se centra en la detección de ojos y fisuras en quesos fraccionados como método de control de calidad del producto terminado. Mientras que la tesis de Wrangborg and Ardö (2007) presenta el uso de imágenes ópticas para obtener información acerca de la formación de rajaduras en quesos cortados. El trabajo de Lukinac et al. (2018) consiste en una revisión del uso de imágenes ópticas en diversos tipos de quesos para el control de aspectos externos tales como color, formación de hongos, cristales de lactato, y de aspectos internos en quesos fraccionados tales como la distribución de ojos, rajaduras y condimentos (hierbas y especias).

Junto a la implementación de los métodos de imagenología antes mencionados, se han estudiado diferentes técnicas de procesamiento de las imágenes obtenidas para extraer la información deseada. En el caso de desear detectar cavidades de gas, dependiendo de la calidad y tipo de imagen obtenida, las alternativas de procesamiento más empleadas son binarización y detección de regiones (Caccamo et al., 2004; Musse et al., 2014), detección de círculos mediante transformada de Hough (Schuetz et al., 2013) y *template matching* (Schuetz et al., 2013; Kraggerud et al., 2009).

3.1.4. Propuesta de la tesis

El desarrollo de técnicas no destructivas para la determinación de los ojos es de gran interés para la industria quesera. Si bien existen trabajos publicados basados en técnicas de imagenología que permiten este objetivo, las mismas son demasiado complejas y tienen un costo muy alto o rango limitado como para su aplicación industrial. Por otra parte, aunque varios autores han mencionado la posible relación entre los cambios geométricos en la horma de queso y la formación de ojos (Huc et al., 2014b,a), no se ha reportado hasta la fecha ningún trabajo que vincule de forma directa estos fenómenos .

En esta tesis se presenta una propuesta no destructiva para estimar la formación de ojos en quesos tipo *Emmental* a partir de su geometría y masa. Esta propuesta se validó estimando el volumen de ojos en quesos mediante el procesamiento de fotografías de quesos cortados longitudinalmente en distintos estados de su maduración. Una vez validada la metodología de estimación del volumen de ojos basada en la geometría y masa de las hormas, el resultado consistió en un método rápido, sencillo y no destructivo que permite evaluar la evolución de los ojos durante la maduración de quesos.

⁰A excepción de los publicados en el marco de esta tesis

3.2. Metodología

En esta sección se presentan los métodos empleados para la adquisición de los datos experimentales y su posterior procesamiento. La experiencia consistió en el monitoreo de la maduración de ocho hormas de queso durante 30 días, tal como se explica en 3.2.1. Las hormas eran medidas y pesadas diariamente como se detalla en 3.2.2, volteándose cada 2-3 días según la práctica usual de maduración. Cada 5 días se interrumpió la maduración de una de las hormas, la cual se cortó longitudinalmente por su eje de axisimetría para fotografiar ambas mitades, tal como se describe en 3.2.4.1. Mediante el procesamiento de las imágenes se pudo obtener información geométrica de los ojos como se presenta en 3.2.4.2 y 3.2.4.3. Dado que en las imágenes se obtiene una proyección de los ojos, se desarrolló un modelo que permite vincular esta proyección con el volumen de los ojos (sección 3.2.5). El objetivo consistió en estimar el volumen de ojos en los quesos durante la maduración usando las imágenes, y emplear este resultado para validar el método de estimación de volumen de ojos basado en las medidas geométricas y la masa de las hormas (ver 3.2.3).

3.2.1. Obtención y conservación de los quesos

Para la obtención de los datos necesarios se realizó un ensayo de maduración con ocho hormas de queso tipo *Emmental*. Las hormas de queso fueron provistas por “Granja Naturalia” y eran quesos que finalizaron la maduración en cámara fría. En el marco de la tesis se estudió la etapa maduración cámara caliente de estos quesos. Las hormas de queso eran cilíndricas, tenían una masa inicial de $8,8 \pm 0,6$ kg, y sus dimensiones eran de $26,2 \pm 0,2$ cm de diámetro y $15,5 \pm 0,8$ cm de altura. Los productos fueron conservados en cámaras refrigeradas en condiciones de temperatura controlada. Las hormas de queso se monitorearon por 30 días, conservándose a $16^\circ\text{C} \pm 1^\circ\text{C}$ y 75 ± 5 % HR (humedad relativa).

3.2.2. Masa, densidad del queso y geometría de las hormas

Durante el proceso de maduración se registraron las medidas geométricas y la masa de cada horma con el fin de estimar el volumen de ojos al combinar

estas medidas con el procesamiento de imágenes. Con el fin mencionado, las hormas de queso se pesaban diariamente empleando una balanza de laboratorio (Ohaus SD 35, OHAUS CORPORATION, EEUU). Tras esta operación, se registraban las medidas geométricas de los quesos según la figura 3.1A. Se registraron tres dimensiones como medidas representativas de este tipo de geometría: el diámetro del cilindro (D), la altura del cilindro (h) y la altura máxima de la horma (H). Estas medidas se determinaron empleando una cinta métrica flexible con apreciación de 1 mm. Para el análisis del comportamiento de las mismas durante la maduración, las medidas se expresaron respecto a las medidas iniciales de cada horma, y luego se calculó el promedio y desviación estándar de las mismas para su estudio a lo largo del tiempo de maduración.

Las hormas se modelaron como cilindros con un casquete esférico en su tapa superior, tal como se muestra en la figura 3.1B. La estimación del volumen de la horma se realizó según la ecuación 3.1, donde D, h, H son las dimensiones geométricas registradas. Este modelado permite una estimación simple del volumen real de la horma, el cual podría determinarse de forma más precisa empleando el desplazamiento de líquido por inmersión o mediante la reconstrucción espacial empleando imágenes de dos cámaras calibradas.

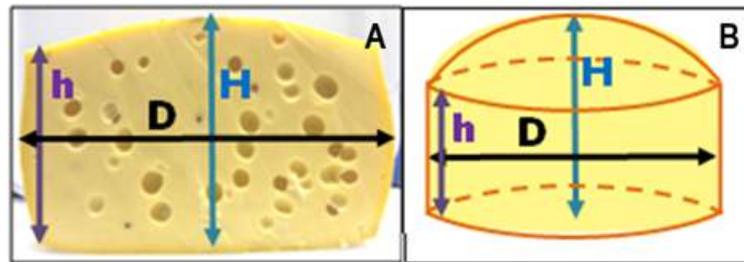


Figura 3.1: (A) Medidas registradas en una horma de queso. (B) Esquema del modelado del volumen de la horma de queso.

$$V_{queso} = \pi h \left(\frac{D}{2}\right)^2 + \frac{\pi}{6} \times (H - h) \times \left[\frac{3D^2}{4} + (H - h)^2\right] \quad (3.1)$$

Ecuación 3.1: Estimación del volumen de la horma de queso.

La densidad de la matriz queso se estimó cada vez que se cortó una horma. De cada horma se extrajeron muestras de 8 ± 2 g de queso sin ojos provenientes de

zonas ubicadas a más de 2 cm de la cáscara, para así asegurar homogeneidad en las propiedades. Con las muestras obtenidas, se estimó la densidad de la matriz queso empleando el método hidrostático de Arquímedes y para cada día de corte se informó un valor promedio de estos resultados. En esta medida se asume que la densidad de la matriz queso es aproximadamente igual en todos los puntos lejanos a la zona de la cáscara, la cual cubre la parte exterior del queso y tiene un espesor despreciable en comparación al tamaño de la horma.

3.2.3. Estimación del volumen de ojos en el queso a partir de la geometría, masa y densidad

Se propone en esta tesis estimar el volumen de ojos dentro de la horma de queso basándose en el volumen y masa de la horma de queso, junto a la densidad de la matriz queso. Se parte de que el volumen de la horma de queso es la suma del volumen de la matriz queso (V_q) y el volumen de ojos (V_o). Si planteamos la definición de densidad para la matriz queso, ρ_q , y asumimos que la masa de la matriz queso es la masa de la horma M , se obtiene la ecuación 3.2. A partir de esta ecuación, se puede despejar V_o .

$$V_h = \frac{M}{\rho_q} + V_o \quad (3.2)$$

Ecuación 3.2: Relación del volumen de la horma de queso con la masa, densidad de la matriz queso y volumen de ojos en su interior.

En esta tesis, la idea consistió en determinar la masa, densidad de la matriz queso y el volumen de la horma como se menciona en la sección 3.2.2. Para validar esta propuesta, se estimó el volumen de ojos dentro de la horma mediante el procesamiento de las fotografías de los quesos cortados. Tras el procesamiento de estas imágenes, se obtuvo el área y distribución espacial de los ojos (sección 3.2.4). Empleando esta información, se hace una proyección espacial de los ojos mediante un modelo que se describe en la sección 3.2.5. Con el volumen de ojos estimado a partir de las imágenes, la densidad de la matriz queso y el volumen de la horma, se estimó la masa del queso y se comparó con su valor experimental. Se consideró que si hay coincidencia entre la masa experimental y la masa estimada, entonces la propuesta permite estimar

de forma adecuada el volumen de ojos usando la densidad del queso junto a la geometría y masa de las hormas.

3.2.4. Obtención y procesamiento de las imágenes

Con el objetivo de validar la propuesta de estimación del volumen de ojos en quesos presentada en 3.2.3, se decidió estimar el volumen de ojos a partir de la proyección espacial de los ojos que se registraron en fotografías de hormas de quesos cortadas longitudinalmente. Durante el ensayo de maduración efectuado se cortó una horma de queso cada 5 días. Las imágenes obtenidas fueron procesadas digitalmente con el fin de segmentar los ojos, obteniendo su área y ubicación en la horma de queso. Sobre cada imagen se realizaron dos tipos de segmentación: manual (imágenes “estándar”) y semi-automática (imágenes “prueba”). Si bien el conjunto de imágenes obtenido en esta tesis fue lo suficientemente pequeño para realizar la segmentación de ojos de forma manual, se estudiaron alternativas de segmentación semi-automáticas para el caso en que fuese necesario realizar la validación de la propuesta con un conjunto mayor de imágenes. Esta sección describe el registro de las imágenes, su procesamiento y el modelo empleado para la proyección espacial de los ojos registrados en las imágenes.

3.2.4.1. Registro de las imágenes y recursos para su procesamiento

Tal como se mencionó anteriormente, cada 5 días se cortó longitudinalmente una horma de queso, siendo dicho corte efectuado en un plano que contenía el eje de axisimetría. Ambas mitades fueron fotografiadas bajo luz fluorescente con fondo blanco uniforme. Las imágenes se obtuvieron con una cámara digital (SM-J500M, Samsung, Corea del Sur) con resolución de 13 MPx y lente de 28mm, sin flash, con balance de blancos automático, nivel de ISO 160, $f = 1/1.9$ y tiempo de exposición $1/25$ s. La distancia entre la cámara y el queso se mantuvo en 55 ± 10 cm. Las imágenes se registraron en formato jpg, siendo las dimensiones originales de 3200 x 1856 píxeles, con resolución horizontal y vertical de 72 ppp.

Las imágenes obtenidas fueron procesadas usando el software Python 3.7 (Van Rossum and Drake, 2011), empleando las librerías Numpy 1.15.4 (Harris et al., 2020), OpenCV 3.4.1 (Bradski and Kaehler, 2000), Scikit-image 0.14.1 (Van der Walt et al., 2014) y SciPy 1.1.0 (Virtanen et al., 2020). Los algoritmos de procesamiento se ejecutaron en un computador portátil HP Pavilion Laptop 15 (HP, HP Inc, EEUU), con 8 Gb de RAM, procesador AMD A12-9720 RP 2.70 GHz con gráficos RADEON R7 (AMD, Advanced Micro Devices, Inc, EEUU).

3.2.4.2. Procesamiento de las imágenes

La primer operación fue recortar manualmente las imágenes para obtener imágenes en las que el queso ocupara la mayor parte del campo de pixeles, resultando en imágenes de tamaño variado. Luego, se efectuó una reducción a la mitad del tamaño de las imágenes mediante interpolación cúbica con el algoritmo `resize` de OpenCV. Esta reducción de tamaño fue realizada con el objetivo de disminuir el tiempo de cálculo de los algoritmos sin comprometer la calidad de las imágenes. Sobre una copia de las imágenes de tamaño reducido se efectuó la segmentación manual (página 46), denominando a este conjunto imágenes “estándar”. Sobre otra copia de las imágenes de tamaño reducido se ensayaron los algoritmos de detección semi-automática (ver 3.2.4.3), las cuales se llamaron imágenes “prueba”.

Sobre ambos conjuntos se realizó la segmentación de la horma de queso respecto del fondo, filtrado, determinación de los ojos y extracción de información geométrica de los ojos. Las operaciones efectuadas en ambos conjuntos de imágenes fue idéntica, a excepción de la segmentación de los ojos de los quesos. A su vez, los resultados de la segmentación semi-automática se compararon con los de la segmentación manual, la cual se consideró el mejor resultado de segmentación para las imágenes. A continuación, se describe en detalle las operaciones efectuadas en el procesamiento de las imágenes.

Segmentación de la horma de queso y cálculo de su área

El primer paso fue segmentar la horma de queso respecto del fondo. Para ello, se determinó visualmente el canal de los espacios RGB, HSV y CIE $L^*a^*b^*$ en el que se lograba el mejor contraste entre la horma y el fondo. El canal b^* del espacio CIE $L^*a^*b^*$ permitió la mejor segmentación, lo cual es de esperar ya que las imágenes estaban compuestas por un objeto de color amarillo sobre un fondo blanco, los cuales están asociados a los canales b^* y L respectivamente. Se transformaron las imágenes del espacio RGB al espacio CIE $L^*a^*b^*$ mediante la función (`cvtColor` de `OpenCv`) y se obtuvo la imagen en escala de grises correspondiente al canal b^* .

Se aplicó un filtro gaussiano sobre la imagen del canal b^* para suavizar la imagen y remover ruido. El tamaño del filtro y el valor de σ de la función Gaussiana se escogieron para que fueran proporcional al tamaño de la imagen. El resultado del pasaje al canal b^* y el filtrado se muestra en la figura 3.2B.

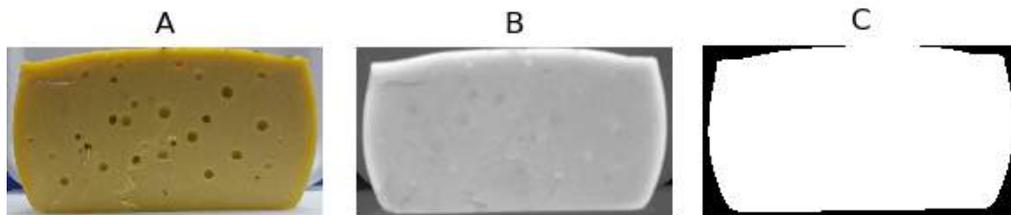


Figura 3.2: (A) Imagen original. (B) Resultado de pasar (A) al canal b^* y su posterior filtrado. (C) Máscara binaria obtenida tras la binarización, detección y relleno de contornos

La imagen filtrada se binarizó con el método de Otsu usando la función `threshold` de `OpenCV`. La imagen binaria resultante consistió en la separación de la horma de queso respecto del fondo. Sobre esta imagen binaria, se realizó una operación de apertura (`morphologicalEx` de `OpenCV`) con el fin de eliminar agrupaciones de píxeles que no correspondan al queso. Se hallaron los contornos de la imagen anterior con la función `findContours` (`OpenCV`) y se escogió el de mayor tamaño, asociado a la horma de queso. A partir de este contorno, se creó una máscara binaria de la horma del queso, tal como se muestra en la figura 3.2C. Se calculó el área de esta máscara mediante suma de sus píxeles y dicha máscara se empleó para eliminar el fondo de las imágenes “prueba”.

Imágenes “estándar”: creación y segmentación

Las imágenes denominadas “estándar” se obtuvieron tras marcar los ojos de los quesos como regiones de color negro (figura 3.3A), editando las imágenes de tamaño reducido según 3.2 con un programa de edición de imágenes (Paint 3D, Microsoft, Windows). Se consideró este conjunto como el mejor resultado de detección de ojos para las imágenes registradas y su segmentación fue el patrón contra el que se contrastaron los algoritmos evaluados.

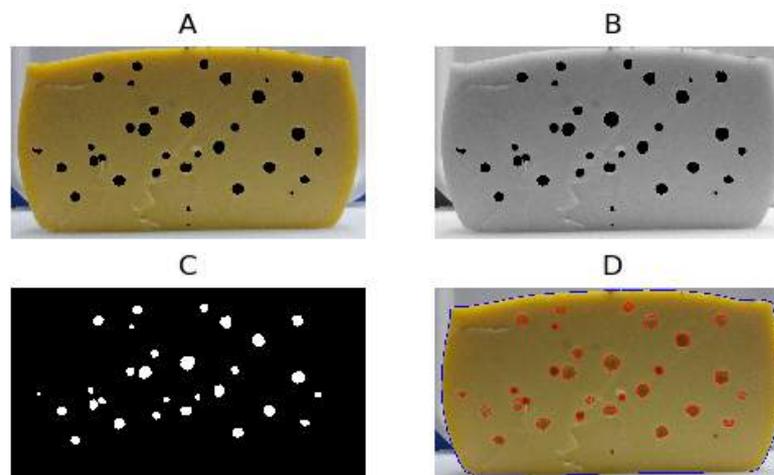


Figura 3.3: (A) Imagen “estándar”, (B) Resultado de pasar (A) al canal L^* y su posterior filtrado. (C) Resultado de segmentación de (B) y posterior aplicación de operadores morfológicos. (D) Resultado de representar los contornos tras etiquetar (C) sobre las imágenes “prueba”.

La segmentación de los ojos en las imágenes “estándar” implicó transformar la imagen del espacio RGB al espacio CIE $L^*a^*b^*$ con la función `cvtColor` (OpenCV) y obtener el canal L^* . La figura 3.3B muestra la imagen del canal L^* en escala de grises, imagen en la cual los ojos son regiones de bajo nivel de intensidad. Sobre esta imagen se realizó una binarización con la función `threshold` (OpenCV) con un umbral tal que permite la segmentación de los ojos marcados. Sobre el resultado de segmentación se aplicaron operaciones de apertura y clausura con la función `morphologyEx` (OpenCV) para eliminar agrupaciones de píxeles aislados y cerrar regiones. El resultado de la binarización y aplicación de operadores morfológicos se muestra en la figura 3.3C.

Sobre la imagen binaria obtenida se etiquetaron las regiones asociadas a los ojos empleando el algoritmo `label` (Ndimimage-Scipy) y se extrajeron sus propiedades geométricas mediante la función `regionprops` (Scikit-image). Mediante este procesamiento se obtiene el número total de ojos, la ubicación de sus centroides, sus áreas, excentricidades y otros parámetros geométricos de relevancia. Finalmente, se detectaron y dibujaron los contornos de la imagen binaria sobre las imágenes “prueba”, usando las funciones `findContours` y `drawContours` (OpenCV) para facilitar la visualización del resultado final. Esto se muestra en la figura 3.3D.

Imágenes “prueba”: Tratamiento previo a la segmentación

Tras obtener la máscara binaria que permite tanto el cálculo del área de la horma junto a su separación del fondo (página 44), el siguiente paso fue segmentar los ojos respecto del queso. Primero, se visualizaron los canales de los espacios RGB, HSV y CIE $L^*a^*b^*$ para obtener el canal con el mejor contraste entre los ojos y el queso. Se escogió el canal de saturación del espacio HSV, canal en el cual los ojos aparecen como regiones brillantes. Se transformó la imagen del espacio RGB al espacio HSV con la función `cvtColor` (OpenCV) y se seleccionó la imagen en escala de grises correspondiente al canal de saturación.

Sobre la imagen en el canal de saturación se aplicó un filtro gaussiano idéntico al usado en la segmentación de la horma de queso (página 44). Se procedió con una ecualización de histograma mediante `equalizeHist` (OpenCV) para mejorar el contraste de la imagen. Sobre la imagen ecualizada, se aplicó la máscara de segmentación de la horma. Tras esta secuencia de operaciones, se obtuvo una imagen como la representada en la figura 3.4B. Esta imagen se suavizó mediante el funcional de Mumford-Shah (Mumford and Shah, 1989), empleando la implementación propuesta por Ambrosio-Tortorelli (Ambrosio and Tortorelli, 1990). En la implementación usada en la tesis, se obtiene la imagen suavizada (figura 3.4C) y la imagen del campo de gradientes de la imagen suavizada. Sobre la imagen de los gradientes se realizó una binarización con el método de Otsu (`threshold`, OpenCV) y el resultado se muestra en la figura 3.4D. Posteriormente, se realizó una suma ponderada de la imagen suavizada obteni-

da con el funcional de Mumford-Shah con la imagen de la horma segmentada en el canal de saturación, siendo los pesos de cada imagen 0,25 y 0,75, respectivamente. Finalmente, se realizó una nueva suma ponderada del resultado anterior con la imagen binaria de los gradientes, correspondiendo a 0,90 y 0,10 los pesos de cada imagen, respectivamente. El objetivo de combinar la imagen suavizada, la del canal de saturación y la de los gradientes fue lograr una imagen con bajo ruido, alto contraste de los ojos y bordes resaltados, de forma de facilitar la segmentación de los ojos. La figura 3.4E muestra el resultado del tratamiento para la mejora de las imágenes sobre las que se evaluarán diferentes algoritmos de segmentación semi-automática.

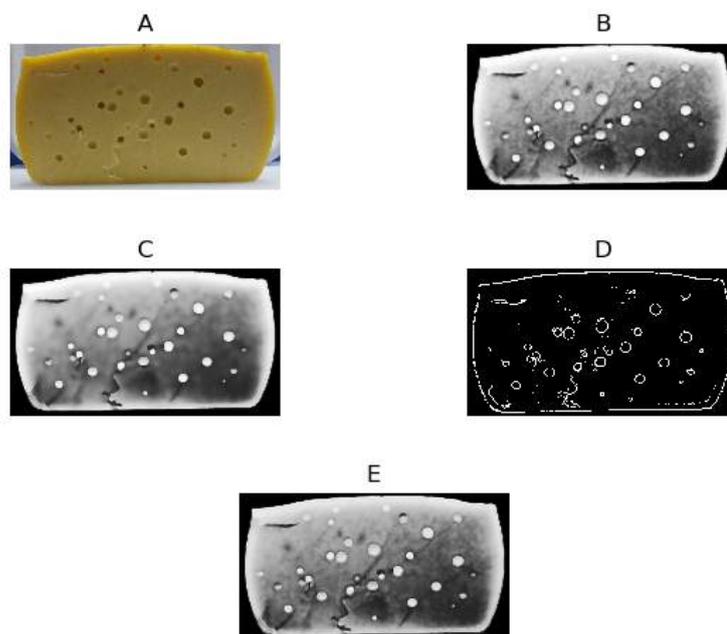


Figura 3.4: (A) Imagen “prueba”, (B) Resultado de pasar (A) al canal de saturación, su posterior filtrado y aplicación de la máscara de segmentación. (C) Imagen (B) suavizada mediante el operador de Mumford-Shah. (D) Gradiente binarizado de aplicar el operador de Mumford-Shah sobre (B). (E) Resultado de sumar (B),(C), (D) de forma ponderada.

3.2.4.3. Algoritmos de segmentación semi-automática y extracción de características

Sobre las imágenes “prueba” obtenidas tras la mejora de contraste descrita, se evalúan tres algoritmos para la detección de ojos:

- (a) Detección de círculos mediante la transformada de Hough circular
- (b) Binarización seguida de detección de contornos
- (c) Detección de blobs

En la detección de ojos empleando (a), el algoritmo usado fue el `circularHough` (OpenCV). `circularHough` tiene parámetros de entrada relacionados a la detección de círculos (distancia entre centros, radio mínimo y máximo, valor de intensidad) y a los umbrales para realizar la detección de bordes de Canny. En (a), el algoritmo primero realiza la detección de bordes de Canny y luego aplica la detección de círculos empleando la transformada de Hough circular. El resultado son puntos con las coordenadas del centro de cada círculo detectado y su correspondiente radio, los cuales se asociaron a los ojos en la imagen del queso.

El algoritmo (b) consistió en primer lugar, una binarización con la función `threshold` (OpenCV). Sobre la imagen binaria se efectuaron una operación de apertura, tres erosiones y tres dilataciones con las funciones `morphologyEx`, `erode` y `dilate` (OpenCV) respectivamente, con el fin de eliminar agrupaciones de píxeles que no se correspondan a ojos. Tras las operaciones morfológicas se realizó la detección de bordes con la función `findContours` (OpenCV). Los bordes detectados se modelaron como elipses con la función `fitEllipse` (OpenCV), determinándose así las coordenadas de sus centros y sus ejes mayor y menor.

El algoritmo (c) consistió en un detector de regiones que se corresponde al algoritmo `SimpleBlobDetector` (OpenCV), el cual tiene como parámetros la intensidad media de la región, el área mínima y máxima, y la convexidad. El resultado fueron anotaciones (*keypoints*) sobre la imagen, conteniendo dichos *keypoints* las coordenadas de las regiones candidatas y el radio del círculo que las describe.

El resultado de aplicar estos algoritmos consistió en una serie de puntos que contienen las coordenadas de los centroides de los ojos detectados en el queso, y en sus radios, en el caso de los algoritmos (a) y (c), o ejes mayor y menor para el algoritmo (b). Tras la aplicación de los algoritmos, los puntos candidatos a clasificarse como ojos fueron filtrados según criterios de forma, área e

intensidad. Se consideraron como ojos las regiones segmentadas que cumplían determinados criterios de área, intensidad y forma. Las regiones candidatas debían cumplir los tres criterios a la vez, de forma de descartar falsos positivos. Se indicó que las regiones que constituían ojos debían tener un área mínima y máxima, las cuales se corresponden a 0,2% y 4,5% del área de la horma de queso, respectivamente. A su vez, dado que en el canal de saturación los ojos tienen valores de intensidad altos, se indicó que las regiones candidatas a ojos debían tener valor medio de intensidad superior a 128. Respecto al criterio de forma, se indicó que la relación entre eje mayor y eje menor de las elipses que describen las regiones segmentadas debe ser menor a 2,5. Por otro lado, se registró el tiempo de cómputo para el procesamiento de cada imagen por cada algoritmo.

3.2.4.4. Evaluación de los algoritmos de segmentación de ojos

Tras la aplicación de los algoritmos de segmentación estudiados (ver página 48), se contrastaron sus resultados con los obtenidos para la segmentación de las imágenes “estándar”. Cuando se posee una imagen etiquetada con la segmentación correcta y se evalúan algoritmos de segmentación para esa imagen, un parámetro comúnmente empleado es el cociente de intersección de las regiones segmentadas de ambas imágenes sobre la unión de las mismas (IoU). Este indicador puede tener valores entre 0 y 1, siendo valores a 0,5 considerados aceptables, 0,7 muy buenos y 0,9 una segmentación casi perfecta.

Por otro lado, para evaluar en mayor detalle el resultado de los algoritmos, se propone comparar la fracción de ojos correctamente ubicados ($FOCU$) y el área total de ojos detectados (A), en todos los casos referidos a los valores de los “estándares”. Se considera que un ojo segmentado por un algoritmo está correctamente ubicado si su centroide y el centroide de un ojo en el “estándar” no difieren en más de 10 píxeles. $FOCU$ es un parámetro propuesto en esta tesis para evaluar la capacidad de los algoritmos de localizar los ojos de forma correcta. $FOCU$ se definió como el cociente de ojos correctamente ubicados (OCU_i) respecto a los ojos del estándar (O_e) ponderado por el cociente de ojos correctamente ubicados sobre el total de ojos detectados por el algoritmo (O_i), según se muestra en la ecuación 3.3.

$$FOCU = \frac{OCU_i}{O_e} \frac{OCU_i}{O_i} \quad (3.3)$$

Ecuación 3.3: Cálculo de la fracción de ojos correctamente ubicados (FOCU).

3.2.5. Estimación del volumen de ojos de quesos a partir de imágenes

Mediante las imágenes “estándar”, el área de los ojos se calculó empleando binarización, etiquetado de las regiones y extracción de las áreas de las regiones etiquetadas, tal como se mencionó en la sección 3.2.4.2. Al trabajar con las imágenes de dos mitades de un plano que contiene el eje de axisimetría de las hormas, se tiene una única proyección de la distribución de los ojos en el interior de las mismas. Sin embargo, se propuso estimar el volumen de los ojos a partir del área registrada en las imágenes. Para ello se asumieron dos hipótesis:

- (I) Las hormas de queso son axisimétricas
- (II) No hay ángulos preferenciales para la formación de ojos en los planos que contienen el eje central de simetría (Huc et al., 2014b)

Si las hipótesis planteadas son correctas, cualquier corte realizado en el plano que contiene el eje de axisimetría es representativo de la distribución interna de los ojos en el queso. Al suponer que todas las secciones de corte en este plano son equivalentes, una forma de reconstruir el volumen consiste en efectuar una rotación de 180° del plano de corte sobre el eje de simetría. De esta forma, la reconstrucción del volumen de los ojos se corresponde con medio toroide (ecuación 3.4). En la ecuación 3.4, A_i es el área de del píxel perteneciente a una región segmentada como ojo, r_i son las coordenadas del píxel referidas al eje central del queso y FE es un factor de escala que permite la conversión de píxeles a m^3 . Por lo tanto, conociendo el área de cada ojo en las imágenes, su distancia al centro de axisimetría y la relación entre el área de un píxel y las dimensiones del mundo real, se puede obtener el volumen que ocupan los ojos dentro de la horma de queso. Este modelo se denominó “modelo de los toroides”.

$$V_{ojos} = FE \times \pi \times \sum_{i=1}^N (A_i \times r_i) \quad (3.4)$$

Ecuación 3.4: Estimación del volumen de los ojos en el queso a partir de imágenes.

Tal como se mencionó en la sección 3.2.3, se estimó la masa de queso como método de validación del “modelo de los toroides”, modelo que permite la estimación del volumen de los ojos en el queso a partir de imágenes. La ecuación 3.5 es la misma que la ecuación 3.2, expresada de una forma que permite obtener la masa del queso a partir de la densidad de la matriz queso (ρ_q), el volumen de la horma de queso (V_h) y el volumen de los ojos (V_{ojos}), el cual fue obtenido mediante el “modelo de los toroides”. La concordancia entre la estimación de la masa y su valor experimental permite decir que la estimación del volumen de ojos a partir de las imágenes junto a la propuesta de la sección 3.2.3 son buenos estimativos del volumen de ojos.

$$M_{est} = \rho_q \times (V_h - V_{ojos}) \quad (3.5)$$

Ecuación 3.5: Estimación de la masa de queso a partir del volumen y densidad de la matriz queso.

3.3. Resultados y discusión

En esta sección se presentan los resultados obtenidos de aplicar la metodología descrita en 3.2 para la determinación del volumen de ojos en quesos tipo *Emmental*. La sección se divide en dos grandes bloques: los resultados del análisis de imágenes y los resultados de las medidas externas del queso. La sección comienza con la presentación de una secuencia de imágenes obtenidas durante la maduración, y a partir de los resultados de la segmentación de las imágenes “estándar” se discute la evolución geométrica de los ojos durante la maduración. Posteriormente, se presentan los resultados de aplicar los algoritmos de segmentación (ver página 48) sobre las imágenes “prueba”, comparando su desempeño contra los resultados de las imágenes “estándar”. Se prosigue con la presentación de los resultados de la masa y la geometría del queso durante la maduración y a partir de estos resultados se realizó la validación de las propuestas para estimar el volumen de ojos en quesos tipo *Emmental*. Esta validación se realizó estimando la masa de la horma del queso empleando el volumen de ojos obtenido con el “modelo de los toroides”, la densidad del queso y volumen de la horma. Tras esta validación, se exploró la relación del volumen de ojos con el tiempo de maduración y la altura de la horma de queso. La sección concluye comentando la relevancia de los resultados obtenidos para el monitoreo de la maduración de quesos formadores de ojos.

3.3.1. Evolución de los ojos en los quesos y resultados de las imágenes “estándar”

En la figura 3.5 se muestran algunas de las imágenes registradas a lo largo de la maduración. Las imágenes muestran los resultados de la segmentación de las imágenes “estándar”, resultados contra los que se contrastaron los algoritmos de detección de ojos. En las imágenes de la figura 3.5, se muestra en línea azul la detección de la horma de queso, y en línea roja, la segmentación de los ojos. Como se mencionó en la sección 3.2.4.2, la segmentación de la horma se realizó de forma automática y es el mismo procesamiento para los conjuntos “prueba” y “estándar”. En cambio, la segmentación de los ojos se realizó de forma manual para conjunto de imágenes “estándar”.

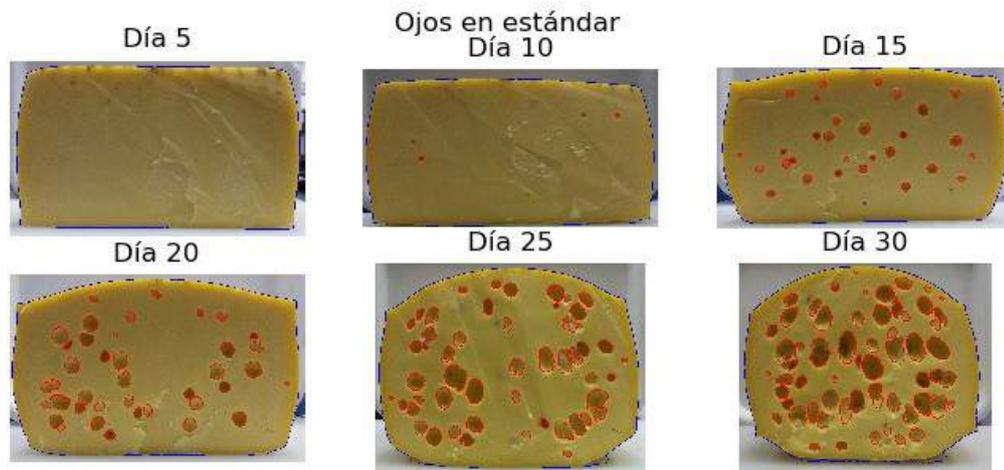


Figura 3.5: Fotografías de las secciones de queso en diferentes días de su maduración. Los contornos en rojo se corresponden con la segmentación de las imágenes “estándar”.

En la figura 3.5 se aprecian dos grandes fenómenos: el aumento del número y tamaño de los ojos, y la deformación de la horma de queso. La deformación de la horma de queso se puede vincular a la formación de los ojos dentro de la horma (página 65). Respecto a la evolución de ojos, en las imágenes de la figura 3.5 se visualiza que los ojos son perceptibles a partir del día 10 de maduración, considerándose que no hay formación de ojos previo a ese momento en las condiciones estudiadas. A medida que avanza el proceso de maduración, aumentan el número, tamaño y excentricidad de los ojos dentro de la horma.

Los resultados de los parámetros geométricos de los ojos, calculados tal como se describe en 3.2.4.2, se presentan en la figura 3.6. La figura 3.6A muestra el área de los ojos respecto al área de la sección fotografiada del queso en las imágenes “estándar” para cada día de maduración, mientras que la figura 3.6B muestra el aumento de excentricidad promedio de los ojos. El aumento del área de los ojos relativa al área de la sección de queso está ligado al aumento de volumen de los ojos para este tipo de queso, consecuencia de los procesos fermentativos que ocurren en el queso. En la bibliografía se reportan diversos métodos para medir este volumen, obteniendo en todos los casos un comportamiento similar al de la figura 3.6A.

El cambio en la excentricidad de los ojos se debe a varios factores. Uno de los factores que explica este fenómeno es la fusión de los ojos, lo que provoca que

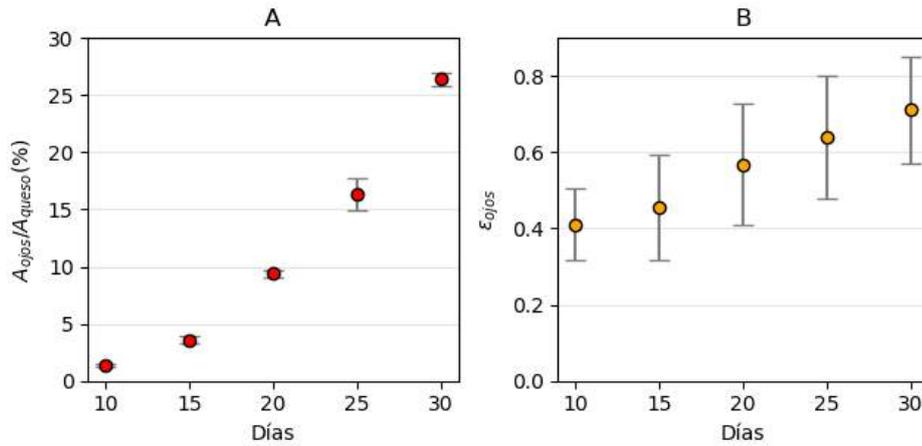


Figura 3.6: Evolución durante la maduración en las imágenes “estándar” del área de ojos respecto al área de la sección de queso (A) y de la excentricidad promedio de los ojos (B). Barras de error se corresponden a la desviación estándar.

la unión de ojos esféricos derive en un único ojo de forma elipsoidal (Schuetz et al., 2013). Otro factor que explica el aumento de excentricidad es que los ojos tienden a expandirse más según el eje vertical que según el eje horizontal. Este fenómeno se debe a la estructura de los gránulos de queso, los cuales se orientan horizontalmente como consecuencia de la etapa de prensado de la horma. A su vez, estos gránulos poseen gran elasticidad ante la presión que generan las burbujas de gas (Huc et al., 2013). Por lo tanto, si bien al inicio de la maduración la formación de una burbuja de gas adquiere forma esférica, al aumentar su volumen se favorece la expansión según el eje vertical debido a la elasticidad de los gránulos de queso en esa dirección.

3.3.2. Resultado de la detección semi-automática de ojos

Tras visualizar la evolución de los ojos en las imágenes “estándar”, se procedió a aplicar los algoritmos de detección de ojos en las imágenes “prueba”. En la figura 3.7 se visualizan la segmentación de ojos lograda con los algoritmos estudiados en quesos en distintos estados de maduración. En la fila superior se muestra el resultado “estándar”, y desde la segunda a la última fila se presentan los resultados del uso de la transformada de Hough circular (a), binarización seguida de detección de contornos (b) y detección de blobs

(c), respectivamente. Contrastando los resultados de los algoritmos con las imágenes “estándar” se puede apreciar que el desempeño de los algoritmos es función del estado de maduración del queso y que en todos los casos hubo falsos positivos y ojos no detectados.

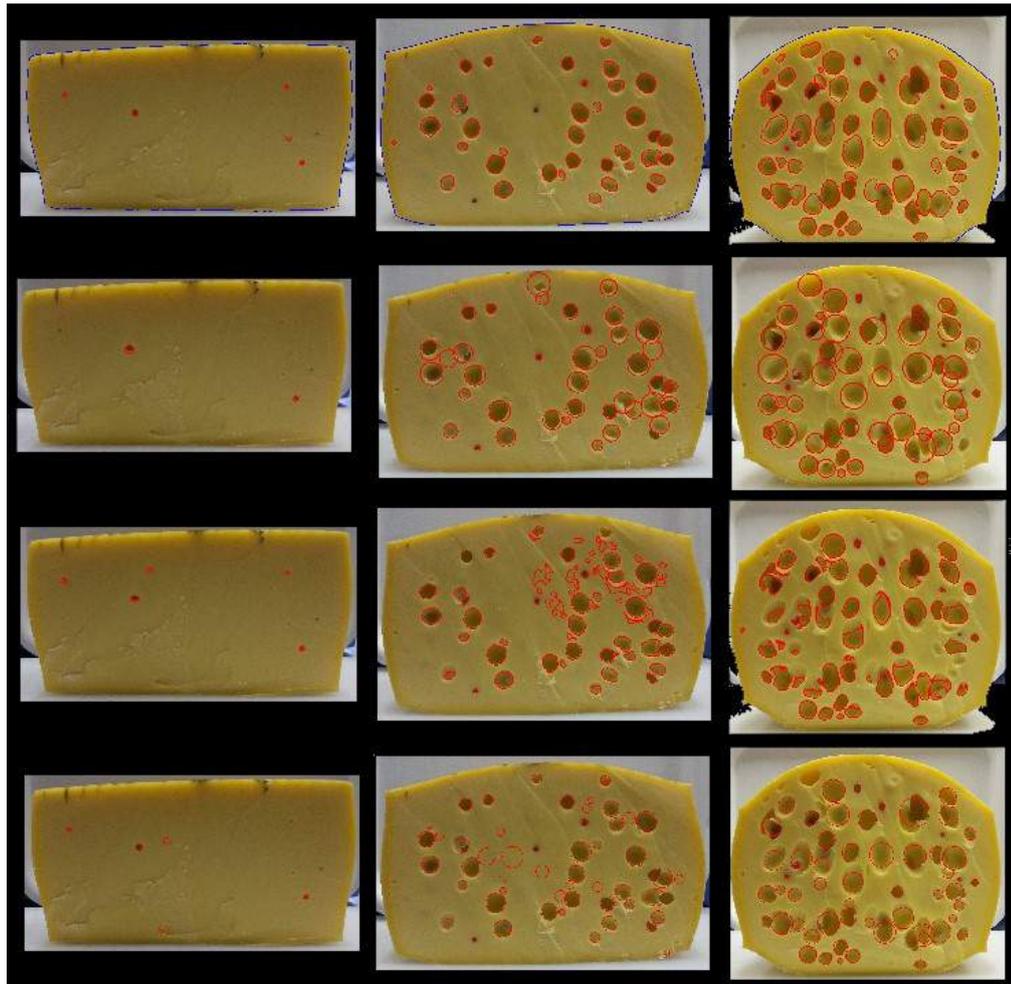


Figura 3.7: Ejemplos de los resultados de segmentación para quesos en diferentes estados de maduración. Columnas (de izquierda a derecha): resultados en los días 10, 20 y 30 de maduración. Filas (de arriba hacia abajo): segmentación del “estándar”, algoritmo (a), algoritmo (b) y algoritmo (c).

En contraste a la visualización de la detección de ojos que se presenta en la figura 3.7, la figura 3.8 presenta indicadores numéricos que permiten evaluar los resultados de los algoritmos de segmentación de forma cuantitativa. Los parámetros escogidos (IoU , $FOCU$ y tiempo de cómputo) se presentan en función del número de ojos en las imágenes “estándar”. La figura 3.8A muestra

el IoU (intersección sobre unión) de la segmentación. Mientras que la figura 3.8B presenta el $FOCU$, el cual es el cociente entre los ojos correctamente localizados respecto al “estándar” ponderado por los ojos correctamente localizados respecto todos los ojos detectados por el algoritmo. Finalmente, en la figura 3.8C se presenta el tiempo de cómputo por imagen t_c .

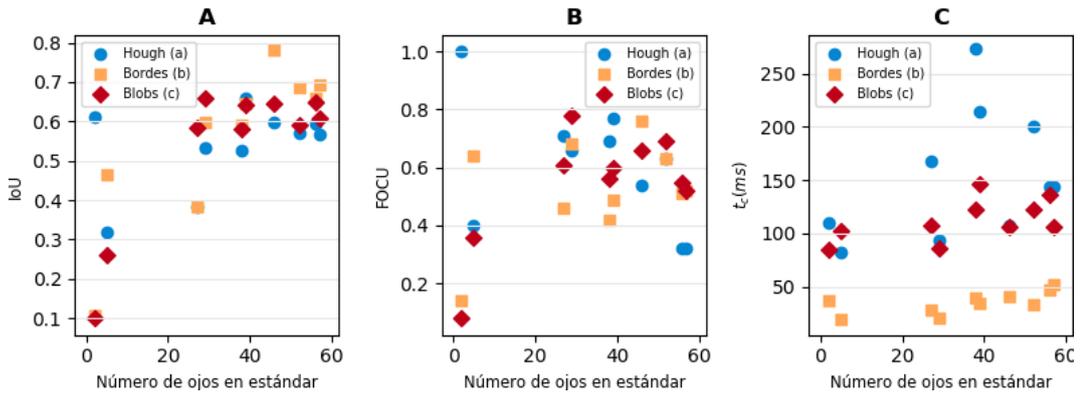


Figura 3.8: Comparación de aspectos de desempeño de los algoritmos en función del número de ojos presente en las imágenes “estándar”. (A) Intersección sobre unión (IoU) de la segmentación de los algoritmos respecto a los “estándar”, (B) fracción de ojos correctamente ubicados ($FOCU$) respecto al total en el “estándar”, (C) tiempo de cómputo (t_c) por imagen en ms.

Los valores de IoU promedio para todos los algoritmos fueron superiores a 0,50 y los de $FOCU$ mayores al 0,52. En particular, los mejores valores de estos indicadores se obtuvieron cuando el número de ojos en el “estándar” fue mayor a 25. La dependencia del IoU y del $FOCU$ respecto al número de ojos en las imágenes “estándar” era de esperar ya que el número de ojos se vincula con complejidad del problema de segmentación. A su vez, ya que los parámetros de los algoritmos fueron escogidos para detectar ojos redondos o elipsoidales de cierto tamaño y distribución espacial, desviaciones de este modelo provocan baja performance en la segmentación. Particularmente, cuando hay pocos o ningún ojo en las imágenes, los resultados de segmentación pueden presentar un número considerable de falsos positivos que impactan negativamente en $FOCU$ y en IoU . Por otra parte, cuando hay formación excesiva de ojos, los mismos presentan mayor grado de deformación y tienden a fusionarse, por lo que los algoritmos pueden presentar problemas en la detección y localización de los ojos en el queso. Respecto al comportamiento del t_c , todos los algoritmos tuvieron una dependencia creciente con el número de ojos debido al aumento

en las operaciones de cálculo.

Por otro lado, la figura 3.9 representa el área total de ojos respecto al área de la sección de la horma (como porcentaje) a lo largo de la maduración. En esta figura se contrastan los resultados de cada algoritmo con los obtenidos para las imágenes “estándar”. Se aprecia en que la figura 3.9 el resultado de segmentación de los algoritmos sigue la misma tendencia que el “estándar”. Por lo tanto, se consideró que los tres algoritmos tuvieron un buen resultado en cuanto a la estimación del área de ojos relativa al área de la sección de queso.

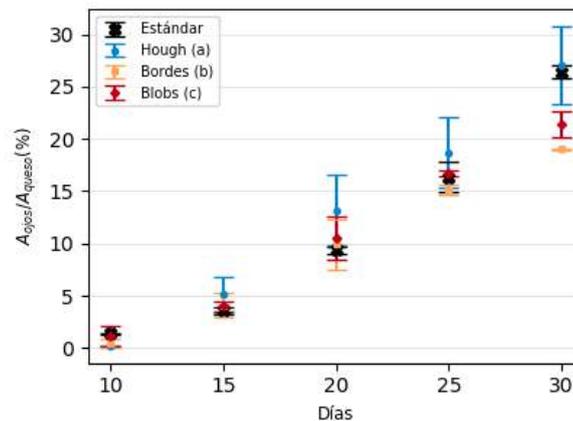


Figura 3.9: Evolución durante la maduración del área de ojos relativa al área de queso (expresada en porcentaje) para la segmentación de las imágenes “estándar” y la segmentación con los tres algoritmos evaluados de las imágenes de “prueba”.

Finalmente, en la tabla 3.1 se presentan el número de ojos detectados, el área de ojos relativa al área de la sección de la horma (en porcentaje) y el número de ojos correctamente localizados tras la aplicación de los algoritmos de segmentación estudiados en esta tesis. Se presentan los resultados de dichos algoritmos para ambas secciones que se obtienen tras cortar las hormas de queso. Las mismas son similares en cuanto son imágenes especulares aunque pueden haber ligeras diferencias debido a la posición de los ojos dentro de la horma respecto al plano de corte.

Analizando los resultados obtenidos para el algoritmo (a), vemos que el área total de ojos hallada tiende a ser mayor que la determinada por el resto de los

Resultados segmentación								
	Sección 1				Sección 2			
Día 5	Estándar	Hough	Bordes	Blobs	Estándar	Hough	Bordes	Blobs
$N^{\circ}ojos$	0	0	3	2	0	0	0	1
OCU		0	0	0		0	0	0
$A_{ojos}/A_{queso}(\%)$	0,00	0,00	0,12	0,12	0,00	0,00	0,00	0,03
$(A_m - A_e)/A_m(\%)$		0,00	0,00	0,00		0,00	0,00	0,00
Día 10	Estándar	Hough	Bordes	Blobs	Estándar	Hough	Bordes	Blobs
$N^{\circ}ojos$	2	2	14	25	5	2	5	5
OCU		2	2	2		2	4	3
$A_{ojos}/A_{queso}(\%)$	0,10	0,15	0,81	1,81	0,22	0,16	0,24	0,45
$(A_m - A_e)/A_m(\%)$		51,1	721,5	1742,7		-27,7	9,9	106,7
Día 15	Estándar	Hough	Bordes	Blobs	Estándar	Hough	Bordes	Blobs
$N^{\circ}ojos$	29	21	27	30	27	21	29	32
OCU		20	23	26		20	19	23
$A_{ojos}/A_{queso}(\%)$	3,80	3,95	3,34	3,69	3,33	6,35	4,94	4,33
$(A_m - A_e)/A_m(\%)$		3,9	-12,0	-2,7		90,6	48,2	30,0
Día 20	Estándar	Hough	Bordes	Blobs	Estándar	Hough	Bordes	Blobs
$N^{\circ}ojos$	39	34	41	41	38	39	53	51
OCU		32	28	31		32	29	33
$A_{ojos}/A_{queso}(\%)$	9,64	10,86	8,21	9,04	9,11	15,57	11,64	11,97
$(A_m - A_e)/A_m(\%)$		12,7	-14,8	-6,1		71,0	27,9	31,4
Día 25	Estándar	Hough	Bordes	Blobs	Estándar	Hough	Bordes	Blobs
$N^{\circ}ojos$	52	44	49	54	46	39	46	55
OCU		38	40	44		31	40	41
$A_{ojos}/A_{queso}(\%)$	15,36	21,06	14,77	16,83	17,34	16,28	15,48	16,49
$(A_m - A_e)/A_m(\%)$		37,1	-3,8	9,6		-6,1	-10,7	-4,9
Día 30	Estándar	Hough	Bordes	Blobs	Estándar	Hough	Bordes	Blobs
$N^{\circ}ojos$	57	43	46	57	56	51	53	57
OCU		28	37	41		30	39	42
$A_{ojos}/A_{queso}(\%)$	25,97	24,49	19,07	20,59	26,83	29,76	18,91	22,28
$(A_m - A_e)/A_m(\%)$		-5,7	-26,6	-20,7		10,9	-29,5	-16,9

Tabla 3.1: Resumen de los valores geométricos de los ojos segmentados con los algoritmos evaluados en las imágenes obtenidas. $N^{\circ}ojos$: Número de ojos detectados, OCU : Ojos correctamente localizados, $A_{ojos}/A_{queso}(\%)$: Relación entre el área de ojos respecto al área de la sección de queso, A_m : Área de ojos computada por segmentación semi-automática, A_e Área de ojos computada por segmentación manual.

algoritmos. Por otra parte, el algoritmo (a) subestimó el número de ojos en comparación a los otros algoritmos. A pesar de ello, el algoritmo (a) obtuvo resultados buenos en todas las etapas de maduración, obteniendo como IoU inferior un valor de 0,32, mientras que su promedio fue de $0,54 \pm 0,11$ (figura 3.8A). Los resultados de este algoritmo pueden explicarse debido a que el mismo busca círculos, y este modelado puede sobreestimar el área de aquellos ojos que son elipsoides. Por otra parte, la detección se realiza sobre un conjunto de imágenes binarias que resultan de aplicar el detector de bordes de Canny. El ajuste de los parámetros de este algoritmo es muy sensible a cambios de contraste suaves como ocurre en algunos ojos, por lo que para detectar la mayor cantidad de ojos posibles se escogieron parámetros que causaron sobreestimación del área total de ojos.

El algoritmo (b), que consiste en una binarización seguida de detección de contornos, obtuvo buenos resultados a pesar de su simplicidad. Este algoritmo obtuvo mejores resultados en las imágenes registradas a partir del día 15 de maduración, alcanzando IoU promedio de $0,63 \pm 0,12$ (Figura 3.8A). Una de las principales desventajas de esta implementación es que la misma tiene un único parámetro para la segmentación de ojos y es el umbral de binarización. Este umbral es dependiente del contraste de los ojos respecto al resto de la escena, por lo que sus resultados están condicionados a la imagen de partida en mayor medida que los otros algoritmos. Por otra parte, tal como se muestra en la figura 3.8C, el algoritmo (b) fue 10 veces más veloz que el resto de las implementaciones estudiadas, obteniendo tiempos de procesamiento de 35 ± 11 ms por imagen.

En cuanto al algoritmo de detección de blobs (c), su performance fue buena en las imágenes obtenidas desde el día 15 de maduración. Considerando los valores de IoU a partir del día 15, el algoritmo (c) obtuvo el mejor IoU promedio de los tres algoritmos ($IoU = 0,63 \pm 0,03$). A pesar de ello, el algoritmo (c) detecta un mayor número de ojos que el resto de los algoritmos y por ende un mayor nivel de falsos positivos. Este fenómeno es más relevante al inicio de la maduración (figura 3.8B). Otra desventaja es que este algoritmo tiende a subestimar el área total de los ojos cuando esta supera el 15% del área de la sección de queso. Este resultado puede deberse a una disminución de la intensidad de los ojos en el canal de saturación sobre el final de la maduración, resultando en blobs detectados de menor tamaño.

3.3.3. Masa y medidas geométricas durante la maduración

En esta subsección se presentan la evolución de la masa y las medidas geométricas de la horma a lo largo de la maduración. Se analizan además cuáles de estas medidas cambian de forma significativa con el tiempo de maduración y los motivos detrás de estos cambios.

3.3.3.1. Evolución de la masa de queso

En la figura 3.10 se muestra la evolución de la masa de las hormas respecto a la masa inicial en función del tiempo de maduración. Tal como se ve en la figura, durante la maduración de los quesos se comprobó la pérdida de masa de las hormas. Este fenómeno era esperado en este proceso y se debe mayoritariamente a la migración de agua desde el interior de la horma hacia el ambiente exterior (Hélias et al., 2007). En quesos que no están recubiertos por películas plásticas o cera, tal como el estudiado en esta tesis, la migración de agua hacia el exterior puede modelarse como un proceso difusivo que sigue una ecuación diferencial de primer orden (Simal et al., 2001). En las condiciones estudiadas, la pérdida de masa representó un 5% de la masa inicial de las hormas en los 30 días de maduración en cámara caliente.

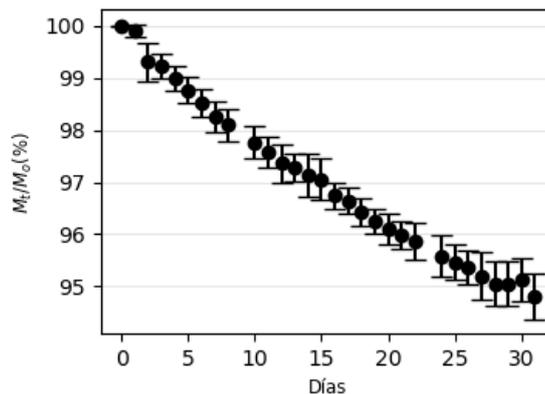


Figura 3.10: Evolución de la masa promedio de las hormas de queso, expresadas como porcentaje de la masa inicial. Barras de error se corresponden a la desviación estándar de la masa de las hormas de queso.

3.3.3.2. Evolución de las medidas geométricas de la horma de queso

Respecto a la evolución de las medidas geométricas externas, se obtuvieron comportamientos diferentes dependiendo de cada medida determinada de acuerdo a las definiciones de la figura 3.1. La figura 3.11A muestra la evolución del diámetro de la horma, la figura 3.11B la altura del cilindro y la figura 3.11C la altura máxima de la horma. Todas las medidas son relativas al estado inicial de cada horma en la maduración.

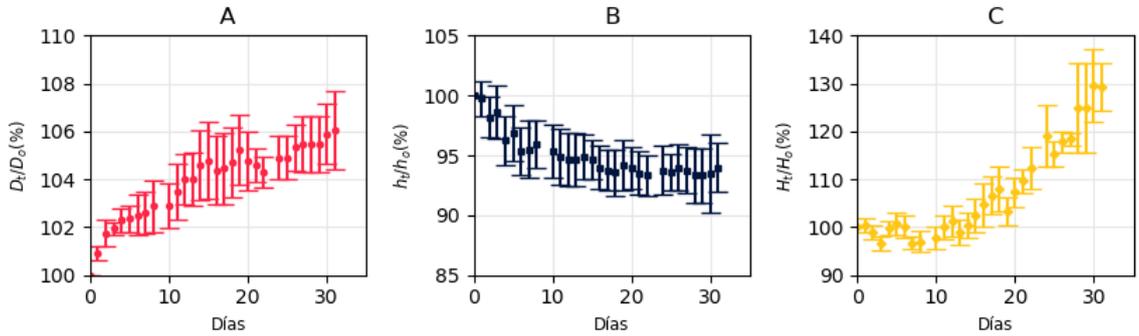


Figura 3.11: Evolución durante la maduración de las medidas geométricas de las hormas de queso. (A) Diámetro de las hormas, (B) altura de los cilindros y (C) altura de los casquetes esféricos, expresados en porcentaje respecto a sus valores iniciales. Barras de error se corresponden a la desviación estándar de las medidas entre las hormas de queso.

La figura 3.11 muestra que tanto la altura máxima relativa como el diámetro relativo aumentaron 30 % y 6 %, respectivamente durante la maduración, siendo el aumento de la altura máxima el más relevante. En cambio, la altura del cilindro disminuyó un 6 %. Estos cambios se pueden explicar por el efecto del peso sobre la matriz viscoelástica del queso. La aplicación de una fuerza normal, tal como el peso, en muestras cilíndricas de queso provoca el efecto de “barreling”, en el cual el diámetro aumenta y la altura disminuye (Fox et al., 2017). Sin embargo, el cambio en la altura máxima de la horma no presenta este efecto, pues aumenta en lugar de disminuir. Esto puede explicarse debido a la formación de los ojos dentro del queso, los cuales aumentaron de tamaño mayoritariamente en la componente vertical (figura 3.5). La formación de los ojos con esta dirección preferencial provoca la expansión de la horma de queso según su eje vertical, causando un aumento de la altura total de la misma. Finalmente, la medida de densidad de la matriz queso no presentó grandes cambios durante la maduración, obteniéndose un valor de $1074 \pm 18 \text{kg/m}^3$.

3.3.4. Validación de las propuestas para la estimación del volumen de ojos

Tras la segmentación de los ojos en las imágenes, el siguiente paso consistió en estimar el volumen de ojos en los quesos empleando el “modelo de los toroides”, como se describe en la sección 3.2.5. El volumen de ojos obtenido a partir

de las imágenes, junto al volumen de la horma de queso y la densidad de la matriz queso se emplearon para estimar la masa de la horma de queso según la ecuación 3.5. Contrastar los resultados de esta ecuación, que deriva de la propuesta presentada en la sección 3.2.3, con los datos de la masa medidos experimentalmente permite validar o descartar la propuesta. El resultado de la estimación de la masa del queso y su contraste con la masa experimental se resume en la tabla 3.2.

Validación del “modelo de los toroides” mediante estimación de la masa			
Día	Masa experimental (kg)	Masa de la ecuación 3.5 (kg)	Diferencia (%)
5	9,00	9,70	7,76
10	7,60	8,20	7,95
15	8,72	9,55	9,54
20	8,96	9,48	5,80
25	8,56	8,26	-3,53
30	7,90	8,21	3,91

Tabla 3.2: Resultados para la validación del “modelo de los toroides”

Los resultados de la tabla 3.2 muestran que la diferencia relativa entre la estimación mediante la ecuación 3.5 y la medida experimental de la masa fue menor al 10 %. Si bien hubo una tendencia a sobreestimar la masa del queso, se puede decir que el “modelo de los toroides” junto a la propuesta descrita en 3.2.3 pueden considerarse como adecuadas para el alcance de este trabajo. Las diferencias entre la masa experimental y la estimada están asociadas tanto al “modelo de los toroides” como a la estimación basada en la geometría y masa de las hormas. Se podrían obtener mejores resultados mediante una mejor reconstrucción del volumen de ojos a partir de las fotografías al emplear un mayor número de cortes de una misma horma. A su vez, se puede mejorar los resultados de la ecuación 3.5 con medidas más precisas del volumen de la horma y la densidad de la matriz queso. Respecto a esto último, el análisis de sensibilidad efectuado indicó que la determinación de la densidad de la matriz queso tuvo la mayor influencia en el error de la estimación de la masa. De todas formas, se considera que el “modelo de los toroides” brinda una buena estimación del volumen de ojos del queso a partir de las imágenes obtenidas. También se considera que la propuesta basada en la geometría y masa de la horma como método no destructivo es un estimador válido del volumen de ojos en el queso.

3.3.5. Evolución del volumen de ojos durante la maduración y vinculación con medidas geométricas

Tras validar el “modelo de los toroides”, se calculó el volumen de ojos relativo al volumen de la horma de queso (V_{ojos}/V_{queso}) empleando los datos del procesamiento de imágenes. El resultado de V_{ojos}/V_{queso} se muestra en la figura 3.12, en la que se observa un comportamiento logístico del cociente V_{ojos}/V_{queso} respecto al tiempo de maduración.

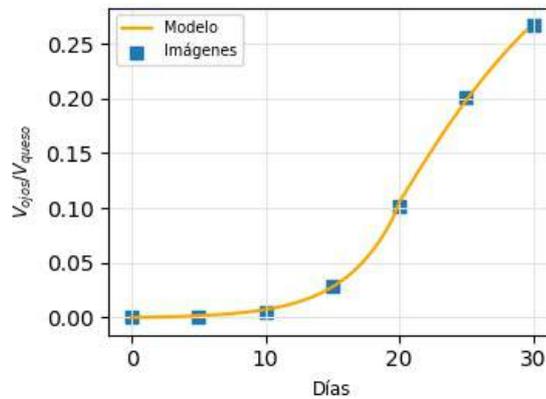


Figura 3.12: Relación entre el cociente V_{ojos}/V_{queso} y el tiempo de maduración. En cuadrados se representan los datos obtenidos con el modelo de los toroides y en línea continua el modelo logístico.

Para tener mayor conocimiento del V_{ojos}/V_{queso} a lo largo de la maduración, se decidió modelar este cociente como un sistema de segundo orden con retardo. Este tipo de modelo permite describir el crecimiento microbiano en sistemas con limitación de sustrato, como es el caso de los microorganismos que hay en el queso. Debido a que la formación de ojos consiste en generación de CO_2 , el cual es un producto de la actividad fermentativa microbiana, se puede aplicar el modelo de crecimiento microbiano para estimar el volumen de CO_2 generado y por tanto el volumen de los ojos en el queso. El modelo se describe en la ecuación 3.6 y los parámetros del ajuste fueron: $K_1 = 5.97 \times 10^{-4}$, $\tau = 3.88$ días, $K_2 = 0.40$, $\tau_1 = 12.37$ días, $\tau_2 = 0.0246$ días, $t_o = 16.33$ días. Cabe mencionar que el error absoluto entre la fracción de volumen de ojos estimado con el “modelo de los toroides” y los valores del ajuste logístico difirieron en menos de 0.005.

$$\frac{V_{ojos}}{V_{queso}} = \begin{cases} K_1 \times (e^{\frac{t}{\tau}} - 1) & t < 20 \text{ días} \\ K_2 \times \left[1 - \frac{\tau_2 e^{-\frac{(t-t_o)}{\tau_2}} - \tau_1 e^{-\frac{(t-t_o)}{\tau_1}}}{\tau_2 - \tau_1} \right] & t > 20 \text{ días} \end{cases} \quad (3.6)$$

Ecuación 3.6: Modelo logístico para el cociente V_{ojos}/V_{queso} en función del tiempo de maduración.

Por otra parte, se vincularon los resultados de las medidas geométricas con el cociente V_{ojos}/V_{queso} para determinar si alguna de estas medidas externas estaba relacionada con la formación de ojos. Se halló que la altura máxima de la horma relativa a la altura inicial se relacionaba con el cociente V_{ojos}/V_{queso} mediante un modelo exponencial de primer orden, tal como se muestra en la figura 3.13. El modelo exponencial que relaciona la altura máxima de la horma y el cociente V_{ojos}/V_{queso} se describe según la ecuación 3.7. En la ecuación 3.7, H es la altura máxima como porcentaje de la altura del primer día, $K = 0,34$, $H_o = 96,4\%$ y $\tau = 22,1\%$. El error relativo entre el modelo propuesto y los datos obtenidos fue menor al 0,01.

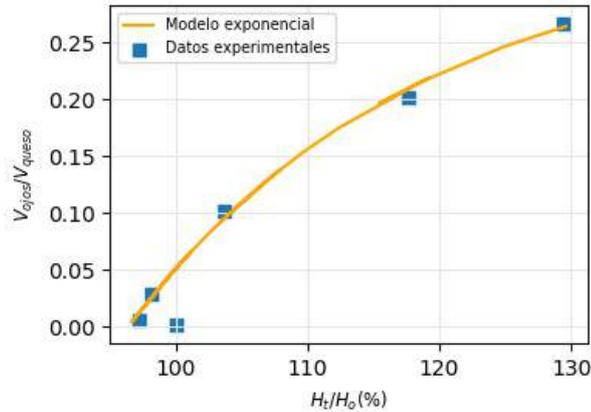


Figura 3.13: Relación entre el cociente V_{ojos}/V_{queso} y la altura máxima de la horma referida a la altura inicial. En cuadrados se representan los datos experimentales y en línea continua el modelo exponencial.

$$\frac{V_{ojos}}{V_{queso}} = K \times \left[1 - \exp^{-\frac{(H-H_o)}{\tau}} \right] \quad (3.7)$$

Ecuación 3.7: Volumen de ojos relativo al volumen de la horma en el tiempo en función de la altura máxima de la horma respecto a la altura inicial.

El resultado obtenido implica que es posible estimar el volumen de los ojos dentro de la horma a partir de un parámetro externo como lo es la altura máxima de la horma. A su vez, la altura máxima de la horma puede determinarse mediante una cámara calibrada y un procesamiento de imágenes como el presentado en esta tesis, ya que es posible obtener una buena segmentación de la horma como se describe en la sección 3.3. Sin embargo, debe tenerse en cuenta que estos resultados fueron obtenidos para el tipo de queso estudiado en las condiciones de maduración y adquisición de datos descritos. Por lo tanto, deben ajustarse los parámetros de los modelos para cada tipo de queso y condiciones de producción particulares. De todas formas, este tipo de estimación puede resultar de utilidad a la industria quesera ya que involucra medidas rápidas, sencillas y no destructivas para la estimación de un parámetro relevante como es el volumen de ojos en el queso. Este parámetro puede vincularse con las condiciones de maduración y puede ser empelado como método de control de dicho proceso.

Capítulo 4

Detección de “manchas pardas” en frutillas

Este capítulo de la tesis se centra en el problema de determinar la presencia de “manchas pardas” en frutillas frescas maduras. Las “manchas pardas” son regiones de color más oscuro que aparecen en la superficie de frutos con mayor grado de envejecimiento. Es por ello que su presencia se asocia a una menor calidad del producto. El objetivo fue conocer si los consumidores son consistentes en la determinación de “manchas pardas” en frutillas frescas, y si este aspecto de calidad puede ser visualizado en un grado similar en fotografías de dichos frutos. En caso de lograr distinguir la presencia de “manchas pardas” en imágenes, el siguiente objetivo fue comprobar si existen diferencias significativas en el color de las zonas afectadas por “manchas pardas” y regiones libres de este defecto en las fotografías de frutillas. Esto último busca conocer la distribución de los valores de intensidad de las “manchas pardas” en determinados espacios de color y así lograr el primer paso hacia la identificación automática de este defecto.

El lector encontrará este capítulo dividido en tres grandes bloques. La primera parte consiste en una reseña de la relevancia de la frutilla como producto agrícola, seguido de la importancia de su aspecto en su calidad junto a una revisión bibliográfica del uso de técnicas de procesamiento de imágenes en la cosecha y poscosecha de frutillas. La primera parte finaliza con una descripción de los objetivos propuestos para el estudio de las “manchas pardas” en frutillas y las hipótesis en las que se basó el trabajo. La segunda parte se centra en la

metodología empleada para validar o descartar las hipótesis propuestas en el primer bloque. En esta segunda parte se describe el entrenamiento de un panel en la evaluación de “manchas pardas”, el registro y extracción de información de fotografías de las frutillas evaluadas y el procesamiento de los datos obtenidos. El último bloque de este capítulo se centra en la discusión y análisis de los resultados obtenidos, estructurado de forma tal que sea consistente con la evaluación de las hipótesis planteadas.

4.1. Frutillas: relevancia, características y estudios previos

La frutilla (*Fragaria x ananassa Duch.*) es una fruta de origen sudamericano que posee una forma, color y sabor característico. Su consumo es mayoritariamente como producto fresco aunque también se emplea como ingrediente en mermeladas, jugos, helados, productos de repostería, etc. Es una fruta de consumo mundial y cuya demanda viene en crecimiento en los últimos años. Según datos obtenidos de la FAO, el promedio cultivado a nivel mundial entre 2015-2018 fue de 8.1 millones de toneladas, siendo los mayores productores China, Estados Unidos, México, Turquía y Egipto.

4.1.1. Frutillas a nivel nacional: producción y preferencia de consumo

En Uruguay la frutilla se cultiva en dos regiones: región norte (Salto y Artigas) y región sur (San José, Canelones y Montevideo). Ambas regiones difieren en los métodos y variedades cultivadas, cubriendo la demanda en distintas épocas del año ([Observatorio Granjero, 2017](#)). La producción es estacional, alcanzando su máximo en octubre y noviembre, con un promedio de 700 ton/mes, y su mínimo entre febrero a mayo, con una producción de 50 ton mensuales. Es un producto de alto valor comercial cuyo precio promedio de ingreso al mercado es de 95 \$U/kg ¹ o 2,3 USD, precio que fluctúa ampliamente dependiendo de la época del año.

¹Precio promedio del periodo 2016-2019

Cuando se presenta para su consumo como producto fresco, su aspecto visual es uno de los principales factores al ser escogido por los consumidores, siendo indeseables aquellos productos inmaduros, de formas subestándar o con avanzado deterioro (Bhat et al., 2015; Cho et al., 2019). En nuestro país, los consumidores prefieren productos de tamaño medio a grande, de color rojo homogéneo y son sensibles a variaciones de precio (Ares et al., 2009; Observatorio Granjero, 2017). A su vez, la frutilla como producto fresco tiene una vida útil limitada por la aparición de defectos superficiales visibles asociados al deterioro del producto, tales como disminución del brillo, manchas y hongos (Matar et al., 2018). Debido a esto, la detección de los defectos visuales del producto puede ser de utilidad en el estudio de su vida útil y auxiliar en la toma de decisiones de conservación del producto. En particular, el uso de herramientas de procesamiento de imágenes es una alternativa rápida que puede emular la evaluación que hace un consumidor a la hora de comprar frutillas.

4.1.2. Estudios previos del uso de imágenes en la evaluación de frutillas

En la literatura se reporta el uso de técnicas de visión computacional y procesamiento de imágenes para diversas operaciones en la producción de frutillas. En particular, se ha hecho énfasis en el uso de estas herramientas en la cosecha de frutillas y la evaluación de sus parámetros de calidad postcosecha. Respecto al uso de procesamiento de imágenes y visión computacional durante la cosecha, los estudios se enfocan en la detección de la frutilla madura (Indrabayu et al., 2019) y su ubicación espacial para su recolección, siendo esta información empleada para la construcción de robots cosechadores (Han et al., 2012; Yu et al., 2019). En esta área de aplicación, se combina la adquisición de fotografías o videos, su procesamiento y tras ello, una acción mecánica por parte del robot.

En cuanto a la evaluación de la calidad de frutillas tras su cosecha, existen dos grandes áreas en las que se aplica el procesamiento de imágenes: calificación de los frutos según su forma, tamaño o color (Oo and Aung, 2018; Liming and Yanchao, 2010; ElMsry et al., 2009), o la detección de defectos como hongos y daño del fruto (Liu et al., 2018). En este tipo de aplicaciones, se emplean

tanto fotografías (Liming and Yanchao, 2010) como imágenes hiperespectrales (Liu et al., 2018) o de infrarrojo cercano (Nagata et al., 2006), brindando cada tipo de imagen información diferente acerca del estado del fruto. Si bien estas nuevas tecnologías permiten inferir información de las propiedades internas del producto al ampliar el espectro electromagnético en el que se registra la información, suelen ser más costosas y complejas en comparación al uso de fotografías.

4.1.3. Propuesta de la tesis

En esta tesis se emplean técnicas de procesamiento de imágenes sobre fotografías de frutillas maduras con el fin de detectar la presencia de manchas oscuras, vulgarmente conocidas como “manchas pardas”, en dichas fotografías. Se trabajó sobre este tipo de defectos debido a que los consumidores los asocian a un producto menos fresco, con cierto grado de deterioro y por ende de menor calidad, pero que aún es comercializable. Empleando esta información, se pueden efectuar operaciones de comercialización para que este producto sea consumido más rápidamente que productos más frescos y así disminuir las pérdidas de este alimento. A su vez, a la fecha de redacción de esta tesis no se ha reportado en bibliografía la detección de “manchas pardas” en frutillas, a diferencia de otros defectos fácilmente distinguibles como la presencia de hongos, daños en la superficie, desviaciones en el color o forma de los frutos. Por otro lado, como se reporta en Matar et al. (2018), la detección de defectos visuales de forma más precisa puede ser de utilidad en la evaluación de diferentes estrategias de conservación de frutillas en comparación a métodos tradicionales de evaluación del deterioro que se basan en el conteo de productos afectados.

Se plantean las siguientes hipótesis:

- (I) Las “manchas pardas” son detectables de forma consistente por los consumidores mediante evaluación visual
- (II) Las “manchas pardas” pueden ser reconocidas en fotografías
- (III) El análisis de las imágenes obtenidas en el punto anterior permite localizar “manchas pardas” en frutillas y estimar su extensión

4.2. Metodología

El objetivo de las operaciones efectuadas en frutillas fue determinar si es posible distinguir la presencia de “manchas pardas” en fotografías de frutillas frescas maduras. Para ello, se entrenó un panel de consumidores en la detección y estimación del área ocupada por “manchas pardas” en frutillas. Este panel realizó la evaluación visual en forma presencial de frutillas frescas. Posteriormente, el panel debió indicar en fotografías de las muestras evaluadas las zonas en las que veían la presencia de “manchas pardas” y estimar la extensión del área marcada. Se procesaron estas imágenes para determinar la consistencia en la estimación de los evaluadores, contrastando los resultados de lo marcado en las imágenes con lo indicado en la evaluación presencial sobre las frutillas frescas. Tras comprobar consistencia del panel en la determinación de este atributo visual en el fruto fresco como en sus imágenes, se construyó una base de imágenes de “manchas pardas” y zonas de frutilla sin manchas para obtener su valor medio de intensidad en tres diferentes espacios de color en los que se pueden representar estos datos. A partir de estos valores, se busca conocer si el color de las muestras permite determinar las regiones de “manchas pardas” en imágenes de frutillas.

4.2.1. Obtención y conservación de las frutillas

Las frutillas fueron adquiridas en mercados locales de Montevideo, Uruguay, durante en el mes de julio de 2020. Se escogieron productos maduros, de tamaño y peso variable y con diversos niveles de extensión de manchas oscuras. Tras la compra, las frutillas se conservaron dentro de bolsas de polipropileno, a $7^{\circ}\text{C} \pm 2^{\circ}\text{C}$. Las frutillas fueron conservadas en estas condiciones durante un máximo de 2 días, desde su compra hasta el momento en que fueron fotografiadas.

4.2.2. Entrenamiento del panel y evaluación visual

Con el fin de evaluar la primer hipótesis de esta tesis (ver página 70) se realizó un entrenamiento y evaluación visual de “manchas pardas” en frutillas maduras. Para ello, voluntarios del Instituto de Ingeniería Química (Facultad

de Ingeniería, UdelaR) se ofrecieron a participar en la evaluación visual de frutillas. Siete evaluadores voluntarios de ambos géneros, con edades entre 25 a 55 años y con distinto grado de experiencia previa en evaluación sensorial participaron de una jornada de entrenamiento y evaluación. En dicha jornada se instruyó al panel en la detección y estimación de la extensión de manchas oscuras en frutillas. Para ello, un lote de frutillas se dividió en dos conjuntos: uno de entrenamiento y referencia, y otro para la evaluación visual. Tras la instrucción del panel en la detección de “manchas pardas” se realizó la evaluación visual, de ahora en más denominada evaluación presencial.

El panel realizó la estimación de la presencia de manchas y su extensión en un conjunto de 16 muestras de frutillas frescas maduras, siendo la evaluación de las muestras de forma individual. Cada muestra estaba en un soporte que la identificaba y que permitía distinguir cuatro planos de su superficie, los cuales también estaban identificados. La evaluación consistió en estimar la extensión de superficie manchada relativa a la superficie total visible de cada uno de los cuatro planos de las muestras. La estimación del área manchada se realizó mediante un sistema de puntajes discreto, en el que cada valor se correspondía con un rango de superficie manchada de cada plano de la muestra. En la evaluación presencial se trabajó con un sistema de diez puntos, en el que cada puntaje se correspondía a incrementos de 10 % en superficie manchada de frutilla.

Dos días después de la evaluación presencial, los integrantes del panel recibieron las imágenes de las muestras evaluadas. Cada imagen mostraba los cuatro planos de cada muestra, concatenados de forma horizontal como se muestra en la figura 4.1. Los evaluadores marcaron sobre las imágenes las zonas que identificaban como “manchas pardas”. A su vez, completaron una planilla con su estimación del área manchada por plano de frutilla, expresado como porcentaje del área del plano visto. Así se obtuvo por cada evaluador 64 estimaciones de área de manchas por conjunto de 16 muestras de frutilla.

4.2.3. Obtención de las fotografías

Previo a la evaluación visual presencial realizada por el panel de evaluadores, se fotografiaron las 16 frutillas escogidas para ser evaluadas. Las muestras fue-



Figura 4.1: Dos ejemplos de las imágenes recibidas por los evaluadores, mostrando frutillas con diferente grado de afectación por manchas pardas. En el caso de (A) el área cubierta por “manchas pardas” es del 17 %, mientras que para (B) es del 55 %.

ron fotografiadas individualmente bajo luz LED en fondo blanco uniforme, el mismo día y en las mismas condiciones en las que se efectuó la evaluación. Se registraron cuatro planos de cada frutilla con el fin de obtener una representación de toda su superficie. Tanto las muestras como los planos fotografiados fueron identificados previamente.

Las imágenes se registraron con una cámara digital a color de 16 MPx (COOLPIX AW100, NIKON Inc., USA). Las fotografías fueron obtenidas sin flash, en modo macro, ISO 125, con balance de blancos calibrado, $f = 1/3.9$, y tiempo de exposición $1/50$ s. La distancia entre la cámara y la frutilla se mantuvo en 17 ± 5 cm. Las imágenes se registraron en formato jpg, con dimensiones de 4608×3456 píxeles con resolución horizontal y vertical de 300 ppp.

4.2.4. Procesamiento de las imágenes

Las imágenes obtenidas fueron procesadas usando el software Python 3.7 (Van Rossum and Drake, 2011), empleando las librerías Numpy 1.15.4 (Harris et al., 2020), OpenCV 3.4.1 (Bradski and Kaehler, 2000), Scikit-image 0.14.1 (Van der Walt et al., 2014) y SciPy 1.1.0 (Virtanen et al., 2020). Dichos programas se ejecutaron en un computador HP Pavilion Laptop 15 (HP, HP Inc, EEUU), con 8 Gb de RAM, procesador AMD A12-9720 RP 2.70 GHz con gráficos RADEON R7 (AMD, Advanced Micro Devices, Inc, EEUU).

El esquema de procesamiento tiene dos etapas: la preparación de las imágenes para enviarlas a los evaluadores y la extracción de información de las imágenes marcadas que se recibieron de los evaluadores. La preparación de las imágenes involucró la concatenación de los cuatro planos de cada frutilla fotografiada, reducción de tamaño de dicha imagen y estimación del área superficial de la frutilla. La extracción de información de las imágenes sobre las que trabajaron los evaluadores consistió en segmentar las zonas marcadas como manchas, el cálculo de su área y la comparación de los resultados de segmentación para cada muestra. A partir de estas imágenes, se extrajeron subimágenes de las zonas con y sin manchas para estudiar si existen diferencias significativas entre ambos tipos de imágenes.

4.2.4.1. Preparación de las imágenes para los evaluadores

Debido a que para cada muestra de frutilla se tenía cuatro fotografías, una por cada plano identificado, se decidió unir esas cuatro imágenes en una única imagen por frutilla. Para ello, se segmentaron las frutillas mediante binarización y se obtuvieron las regiones que contenían cada plano de cada frutilla. En primer lugar, sobre cada imagen registrada se seleccionó el canal a^* del espacio CIE $L^*a^*b^*$ para segmentar el plano de frutilla del resto. Se escogió este canal ya que permite separar la frutilla del su cáliz y de los elementos del fondo. Para obtener la imagen en el canal a^* se transformaron las imágenes del espacio RGB al espacio CIE $L^*a^*b^*$ mediante la función (`cvtColor` de OpenCv) y se obtuvo una imagen en escala de grises correspondiente al canal a^* . Sobre esta imagen se aplicó un filtro gaussiano con tamaño de 15 píxeles y sigma de 5 píxeles. Sobre la imagen filtrada se realizó la binarización con el método de Otsu (`threshold`, OpenCV) para segmentar la frutilla. Al resultado de binarización se le aplicaron tres operaciones de apertura seguidas de tres de clausura, ambas con un filtro de 19 píxeles, usando las variantes de `morphologyEx` (OpenCV). Esta operación se realizó con el fin de remover clústeres de píxeles que no pertenecen a la frutilla y cerrar huecos que hayan quedado de la binarización. Esta imagen binaria se emplea para construir una máscara de segmentación de cada plano de las frutillas.

El siguiente paso fue concatenar de forma horizontal los cuatro planos de la muestra en una imagen única, empleando la máscara obtenida mediante binarización. Para ello, se halla el rectángulo que contiene la región de interés (ROI) y su centroide, siendo cada ROI un plano de la frutilla en la imagen binaria. Con las coordenadas de los cuatro rectángulos de los cuatro planos de cada frutilla, se escogen las dimensiones del mayor rectángulo. Se concatenan los cuatro planos de la frutilla, tomando un rectángulo ubicado en el centroide de cada plano de la frutilla y de dimensiones correspondientes al mayor ROI. Se reduce el tamaño de esta nueva imagen a la mitad mediante interpolación cúbica y finalmente se envía a los evaluadores. Debido a las diferencias en tamaño de las muestras, las imágenes resultantes tuvieron un largo comprendido entre 1888 y 3470 píxeles y ancho entre 404 y 888 píxeles. Por otra parte, se calcula el área de cada plano de cada frutilla sumando los píxeles de las máscaras binarias obtenidas para cada plano. El área total de cada frutilla fotografiada se corresponde a la suma de las áreas de los cuatro planos fotografiados.

4.2.4.2. Extracción de información de las imágenes marcadas por los evaluadores

Tras recibir las imágenes de frutillas con regiones delimitadas consideradas como “manchas pardas” por los evaluadores (ver 4.2.2), el siguiente paso del procesamiento fue extraer la información de estas imágenes. La información a extraer consistió en el área marcada como “mancha parda” en las muestras y la intersección de las “manchas pardas” delimitadas por los evaluadores.

Primero, los píxeles comprendidos dentro de las regiones delimitadas por los evaluadores fueron cambiados por píxeles blancos (255,255,255). Luego, se binarizaron estas imágenes de forma de segmentar las zonas delimitadas y sumando los píxeles segmentados se calculó el área de “manchas pardas” para cada plano de cada frutilla. El área de “manchas pardas” se expresó como porcentaje del área del plano de la frutilla en cuestión. Este procedimiento se realizó para las imágenes enviadas por los siete evaluadores. Finalmente, se calculó el promedio y la desviación estándar de los valores de las áreas para cada plano de cada frutilla, obteniendo 64 valores.

Con el fin de evaluar la segunda hipótesis de la tesis (página 70), las imágenes binarias obtenidas en el paso anterior se compararon entre sí para determinar si había consistencia en las zonas marcadas como “manchas pardas”. Para ello, se sumaron de forma ponderada las siete imágenes binarias de cada muestra de frutilla (una imagen por evaluador) las que contenían las zonas de “manchas pardas” delimitadas por los evaluadores. Este resultado se binarizó con un umbral de 215, correspondiendo este umbral al resultado de la intersección de lo marcado por 6 evaluadores (N-1). Se obtiene una nueva imagen binaria y a partir de ella se calculó el área de la intersección de las “manchas pardas” marcadas por N-1 evaluadores para cada plano de cada frutilla evaluada. Este valor sirve de indicador cuantitativo para determinar si existe consistencia en la ubicación espacial de las manchas de frutillas, tal como se describe en la página 80. A su vez, se evaluó la consistencia en la localización de las “manchas pardas” en las imágenes de frutillas de forma cualitativa, sumando de forma ponderada los resultados de binarización y obteniendo imágenes de mapas de calor con la distribución de zonas marcadas por los evaluadores.

Finalmente, para evaluar la tercer hipótesis de la tesis (página 70) se estudiaron las zonas que se marcaron como “manchas pardas” por la mayoría de los evaluadores. Para ello, se extrajeron subimágenes (patches) de 64 x 64 píxeles de zonas de “manchas pardas” y zonas libres de manchas. El objetivo fue crear una base de imágenes que contengan información de “manchas pardas” y de zonas de frutilla sin manchas. El criterio de selección fue el siguiente: para zonas de manchas debía existir coincidencia en al menos N-1 evaluadores, mientras que las zonas libres de manchas no debían estar marcadas por ningún evaluador. Además, se extrajeron dos patches como máximo de cada plano de frutilla, de forma que los patches sean representativos del conjunto. Con los patches obtenidos, se calculó el valor medio y la desviación estándar de cada canal de color de los espacios RGB, HSV y CIE L*a*b*. Estos valores se graficaron con el fin de visualizar la distribución de los datos en los espacios de color y se registraron para realizar su tratamiento estadístico (ver 4.2.5). Se estudió la intensidad de los píxeles de los patches basándose en la suposición de que los evaluadores se basan en el color de las manchas para su identificación.

4.2.5. Procesamiento de datos

En primer lugar, se evaluó la consistencia del panel en la identificación y estimación de la extensión de las “manchas pardas” en las muestras de frutillas. Para lograr este objetivo, se empleó el coeficiente de Cronbach (Cronbach, 1951). Este coeficiente permite evaluar la consistencia del panel en la evaluación de un mismo elemento. El cálculo del coeficiente de Cronbach se realiza según la ecuación 4.1 y se dice que el panel es consistente si su valor es superior a 0,8 (Nunnally, 1967).

$$\alpha_c = \frac{k}{k-1} \times \left(1 - \frac{\sum_{i=1}^k \sigma_i^2}{\sigma_{X_o}^2}\right) \quad (4.1)$$

Ecuación 4.1: Cálculo del alfa de Cronbach, siendo k el número de elementos a evaluar en cada test, σ_i^2 la varianza de la evaluación X_i , y $\sigma_{X_o}^2$ la varianza de la sumatoria de todas las evaluaciones ($X_o = \sum_{i=1}^k X_i$)

Tras obtener un panel consistente en la determinación de “manchas pardas”, se promedian según el número de evaluadores los puntajes y área de manchas para cada plano de cada frutilla, obteniendo así 64 valores de cada atributo. Con el mismo criterio, se halla la desviación estándar de estos atributos.

Luego, se procedió a evaluar si existen diferencias en las regiones identificadas como “manchas pardas” respecto al resto de la frutilla. Con los valores medios obtenidos de las subimágenes de manchas y no manchas, se realizó un test de ANOVA comparando los valores medios según Tukey (Næs et al., 2010), empleándose un $\alpha = 0,01$. Cuando las diferencias son significativas, se comparan los valores medios de los patches de zonas manchadas y no manchadas en cada canal de cada espacio de color evaluado. Para efectuar los tests estadísticos se empleó el software InfoStat (Grupo InfoStat, FCA, Universidad Nacional de Córdoba, Argentina).

4.3. Resultados y discusión

El análisis de resultados tiene tres etapas que buscan responder las hipótesis planteadas para este producto. La primer etapa se centra en determinar si los evaluadores son consistentes en la determinación de “manchas pardas” en muestras de frutillas y en imágenes de las mismas. En particular, puede ocurrir que el panel sea consistente en la identificación de “manchas pardas” en el producto fresco pero no pueda hacerlo en las imágenes, implicando que para la determinación de las “manchas pardas” se necesita más información que la contenida en la imagen fotográfica. En caso que los evaluadores sean consistentes en la detección de las manchas tanto en el producto fresco como en las imágenes, la segunda etapa consiste en conocer si se pierde información de las “manchas pardas” cuando se registran en las imágenes. Esto implica ver si la estimación presencial coincide con la estimación en las imágenes y evaluar el grado de coincidencia en la delimitación de manchas en las imágenes por parte de los evaluadores. En caso que no haya pérdida de información en el registro de las manchas en las imágenes, la tercer etapa busca conocer si el color es el aspecto visual que determina qué zona es una “mancha parda” en una frutilla. En caso de ser así, se puede emplear esta información para segmentar las zonas de manchas en la frutilla y estimar su extensión y ubicación en el fruto.

4.3.1. Consistencia y capacidad del panel en la detección de “manchas pardas”

Si bien es relativamente fácil detectar visualmente la presencia de “manchas pardas” en frutillas, la determinación de su extensión puede ser compleja, ya que participan múltiples factores que alteran la percepción de un área. En primer lugar, debido al color de las “manchas pardas” puede haber bajo contraste entre las zonas afectadas por este defecto y las zonas libres del mismo. A su vez, la estimación de un área depende de la capacidad del evaluador, las condiciones de iluminación, la distribución de las “manchas pardas” en la frutilla, etc. En este trabajo, se evaluó visualmente la presencia y extensión de “manchas pardas” en cuatro planos de 16 muestras de frutillas, empleando un panel de siete evaluadores entrenados como se describe en la sección [4.2.2](#).

Debido a la variabilidad de las observaciones, el primer paso consistió en evaluar la consistencia del panel en la determinación de las manchas en las frutillas. Para ello se emplearon los cuatro puntajes de cada muestra para calcular el alfa de Crohnbach. Este indicador resultó ser de 0,89 y 0,90 para las estimaciones del área de “manchas pardas” realizadas por el panel en el fruto fresco y en las imágenes, respectivamente. Estos resultados indican que la consistencia del panel fue alta en la estimación de manchas en las muestras de frutillas.

Tras conocer la consistencia del panel en la evaluación de las muestras, se debe conocer la capacidad de estimar la extensión de las “manchas pardas”. Para ello se comparó el valor de área promedio por plano de frutilla que estimó el panel en la evaluación de las imágenes (A_e) y el promedio del valor calculado mediante el procesamiento de dichas imágenes (A_c), tal como se indica en 4.2.4. La figura 4.2 muestra este resultado, y en la misma se aprecia una respuesta lineal en la estimación del área de “manchas pardas” por parte del panel. La regresión lineal de los puntos dio un coeficiente de regresión $r^2 = 0,9655$, con una recta determinada por la ecuación $A_c(\%) = 0,8911 \times A_e(\%) - 2,1961$. Los coeficientes de la recta indican que los evaluadores tienden a subestimar el área marcada en las imágenes en un 10%. Esto se visualiza en la figura 4.2, donde la función identidad (línea punteada) se halla por encima de la regresión lineal de los datos. A pesar de esto, debido a la variabilidad intrínseca del problema, se considera que la estimación del área de “manchas pardas” por parte del panel es buena.

4.3.2. Evaluación del registro de “manchas pardas” en las fotografías

El siguiente paso en el análisis de resultados consistió en determinar si se preserva la información relativa a las “manchas pardas” en las imágenes de las frutillas. Para ello, se compararon los puntajes promedio por plano de frutilla otorgados por los evaluadores en la instancia presencial (P_p) con los puntajes promedio correspondientes a la estimación que realizaron de las áreas marcadas como “manchas pardas” en las imágenes (P_i). La figura 4.3 muestra este resultado, obteniéndose una relación lineal entre ambas estimaciones. La regresión lineal de los puntos sin considerar los outliers (en círculos naranja en la figura

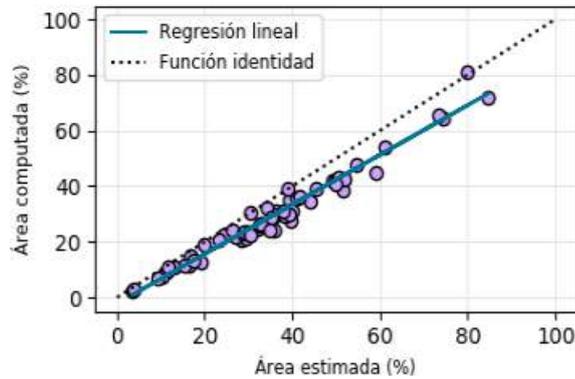


Figura 4.2: Promedio del área de “manchas pardas” por plano de frutilla obtenida mediante procesamiento de imágenes en función del área promedio por plano estimada por los evaluadores en las imágenes marcadas. En línea continua, regresión lineal de los datos. En línea punteada, función identidad.

4.3) dio un coeficiente de regresión $r^2 = 0,7166$, con una recta determinada por la ecuación $P_i = 0,8056 \times P_p + 0,6893$.

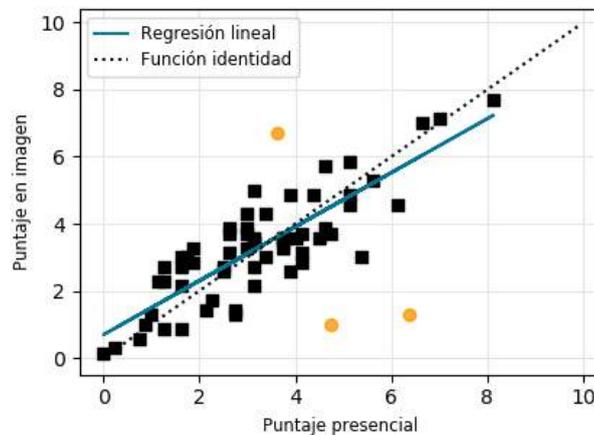


Figura 4.3: Comparación entre el puntaje asignado a “manchas pardas” en imágenes de frutillas en función del puntaje asignado a las mismas muestras en su evaluación visual presencial. En negro, los puntos considerados para el ajuste y en naranja puntos *outliers*. En línea continua, regresión lineal de los datos. En línea punteada, función identidad.

La figura 4.3 nos indica que hay una tendencia lineal en la estimación de la extensión de las manchas. Sin embargo, existen ligeras diferencias en la estimación dependiendo del grado de extensión de las “manchas pardas”, tal como se ve al comparar la regresión de los datos contra la función identidad. Estos

efectos perceptuales pueden deberse a la diferencia en la metodología empleada para la estimación de manchas en las dos instancias. En la estimación de las “manchas pardas” en las imágenes, los evaluadores debían primero delimitar la zona de mancha y luego estimar su extensión. El proceso de delimitar la zona manchada involucra un nivel de estimación más preciso que visualizarlas sobre la muestra y estimar su extensión, que fue el procedimiento efectuado en la evaluación presencial. A su vez, no se descarta que haya cierta pérdida de información en la estimación del área de manchas en las imágenes. Esto puede deberse tanto a la adquisición de las imágenes (cámara e iluminación fundamentalmente) como a la pérdida de profundidad asociada a la obtención de las imágenes. De todas formas, se considera la correlación entre la estimación presencial y la obtenida en las imágenes del área de manchas es lo suficientemente buena para emplear imágenes como método para la detección de este defecto en frutillas.

4.3.2.1. Consistencia en la localización de “manchas pardas” en las imágenes

Si bien la figura 4.3 nos indica que el área de “manchas pardas” en la evaluación presencial y en las fotografías es similar, no nos indica qué zonas están considerando los evaluadores como “manchas pardas”. Para conocer si hay consistencia en la localización de las “manchas pardas” se emplearon las imágenes marcadas por los evaluadores, comparando las áreas marcadas en cada plano de cada frutilla. A partir de estas imágenes, se calculó el área de la intersección entre lo marcado por $N-1$ evaluadores (AI_{N-1}) y el promedio de las áreas marcadas (A_c), en ambos casos por cada plano de frutilla. Los resultados de esta comparación se presentan en la figura 4.4, donde se muestran los resultados para los cuatro planos de todas las imágenes de las muestras.

Los datos de la figura 4.4 se ajustaron a una recta, la cual está dada por la ecuación $AI_{N-1}(\%) = 0,8896 \times A_c(\%) - 10,078$, con un coeficiente de regresión $r^2 = 0,8043$. La relación lineal de la figura 4.4 indica que, si bien los evaluadores marcaron distintas extensiones de “manchas pardas”, en general hubo un consenso en la localización de dichas zonas. El valor de la pendiente, cercano a uno, indica que hay buena coincidencia en las zonas marcadas. El

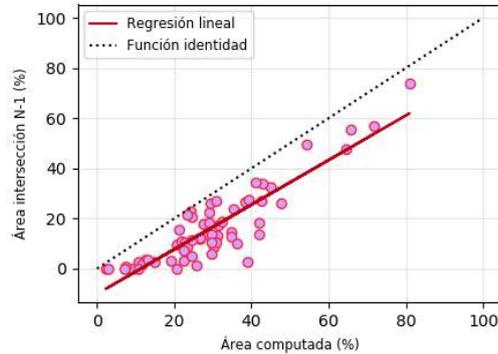


Figura 4.4: Área de la intersección de las “manchas pardas” indicadas por al menos N-1 evaluadores en función del valor de área promedio indicado por todos los evaluadores para cada plano de frutilla considerado. En línea continua, regresión lineal de los datos. En línea punteada, función identidad.

valor de la ordenada es de esperar pues indica que el área marcada promedio es mayor que la intersección de las áreas marcadas por N-1 evaluadores. Cuanto mayor sea la coincidencia de los evaluadores en la determinación espacial de las manchas en las imágenes, la pendiente tendría un valor cercano a uno y la ordenada en el origen sería cero, obteniéndose la función identidad entre ambas estimaciones, como se ejemplifica en línea punteada en la figura 4.4.

La figura 4.5 muestra un ejemplo visual de la consistencia en la determinación de las “manchas pardas” por parte de los evaluadores. En la figura 4.5A se representa el mapa de coincidencias entre evaluadores para la muestra escogida, correspondiendo los valores más cercanos al blanco a un mayor número de evaluadores que marcaron dicha zona. En la figura 4.5B se muestra la conversión de la figura 4.5A en una máscara color verde y aplicada sobre la imagen original. Se observa una tendencia a incluir los bordes de la imagen como “manchas pardas”, lo cual es esperado ya que el contraste con el fondo y los efectos de sombreado pueden inducir a que estas zonas se vean como más oscuras. Sin embargo, este problema se reduce al considerar la intersección de zonas delimitadas. Tal como se ve en la figura 4.5A, los mayores niveles de intensidad no están limitadas exclusivamente a los bordes de la frutilla sino a las zonas donde los evaluadores detectan “manchas pardas”.

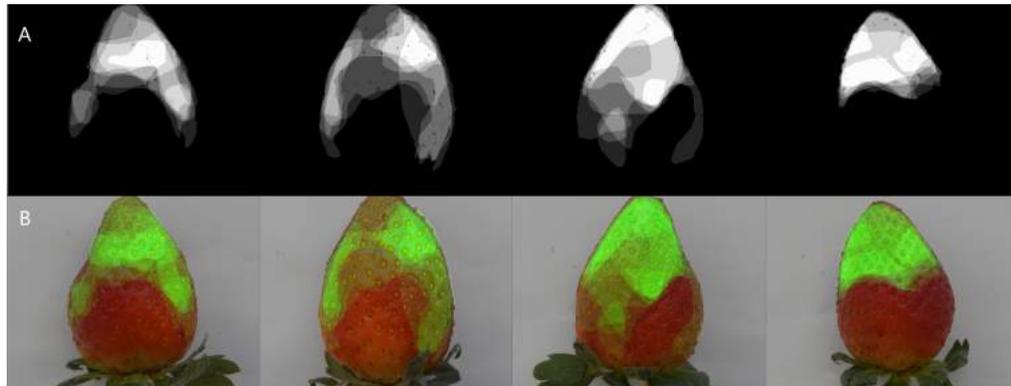


Figura 4.5: Ejemplo de la consistencia en la determinación de zonas de “manchas pardas” sobre las imágenes. (A) Mapa de regiones delimitadas por los evaluadores. (B) Resultado de aplicar (A) con color verde sobre la muestra original.

4.3.3. Análisis de diferencias en el color de zonas con y sin “manchas pardas”

Como se mencionó en la sección 4.2.4.2, tras conocer las zonas que los evaluadores determinaban como “manchas pardas” en las imágenes, se extrajeron subimágenes (patches) de 64 x 64 píxeles con el objetivo de crear una base de patches de “manchas pardas” y zonas libres de “manchas pardas”. Los patches de “manchas pardas” fueron escogidos de las zonas donde hubo coincidencia de al menos N-1 evaluadores en la detección de manchas. En cambio, los patches de zonas sin manchas se escogieron de regiones que ningún evaluador delimitó como mancha. A su vez, no se extrajeron más de dos patches de cada plano de frutilla, con el fin de tener muestras representativas del conjunto. Con estos criterios, se obtuvieron 72 patches de zonas de manchas y 100 patches de zonas sin manchas. Estos patches se transformaron a los espacios RGB, HSV y CIE $L^*a^*b^*$, se extrajeron sus canales y se calculó el valor medio de intensidad de dichos patches en cada canal de color.

Se graficó el resultado de los valores medios de cada canal del conjunto de “manchas pardas” y “no manchas” en ejes cartesianos donde cada eje se correspondía a un canal del espacio correspondiente. Esto se realizó con el fin de visualizar la distribución espacial de estos puntos en los espacios de color RGB, CIE $L^*a^*b^*$ y HSV, representándose los resultados en la figura 4.6.

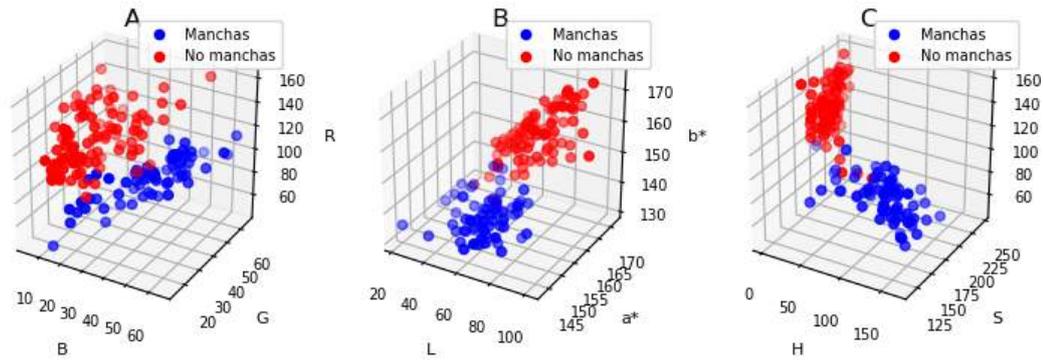


Figura 4.6: Representaciones tridimensionales en los espacios RGB (A), CIE $L^*a^*b^*$ (B) y HSV (C) de los valores medios de intensidad de los patches de zonas con y sin manchas en imágenes de frutillas, obtenidos según 4.2.4.2

Se aprecia en la figura 4.6 que los puntos que pertenecen a “manchas pardas” se agrupan en un clúster diferente al de los puntos sin “manchas pardas”. En particular, existe una mayor separación entre ambos conjuntos de datos en los espacios CIE $L^*a^*b^*$ (figura 4.6B) y HSV (figura 4.6C). Este resultado es de esperar ya que ambos espacios se asocian mejor a la percepción del color que el espacio RGB. En particular, es de interés conocer cual de estos dos espacios permite una mejor separación de los datos. De esta forma, se podría conocer si una zona de una imagen de frutilla es mancha o no empleando el valor medio de intensidad de color representado en un espacio adecuado. Esto permitiría automatizar el proceso de calificación de frutillas según la extensión de su superficie manchada sin la necesidad de un panel de evaluadores.

Se compararon los valores medios de intensidad promedio de los patches en cada canal de los espacios de color, representando estos valores en la figura 4.7. Al contrastar las figuras 4.7A, 4.7B y 4.7C, se aprecia que para el espacio RGB como el espacio CIE $L^*a^*b^*$ hay superposición de las barras de desviación estándar en los promedios de al menos uno de los tres canales. Este fenómeno no ocurre en el caso del espacio HSV, indicando que este espacio sería el que mejor permite distinguir diferencias en los valores de intensidad entre las zonas de “manchas pardas” y las zonas libres de manchas.

Por otro lado, se realizó el análisis de medias empleando el test de ANOVA. Se compararon los valores de intensidad medios de los patches mediante el análisis de medias según Tukey (ver sección 4.2.5). Los resultados para un valor de

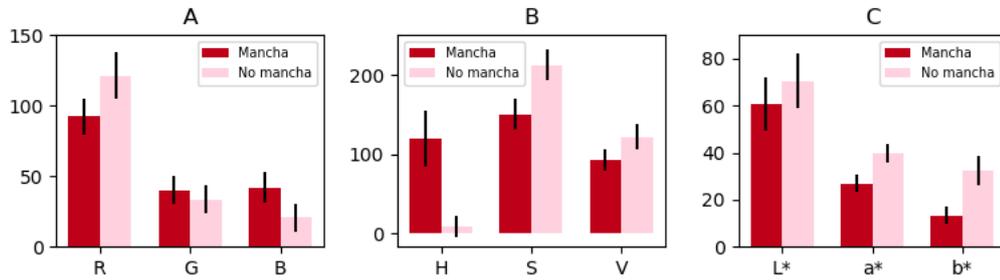


Figura 4.7: Valores promedio de la intensidad media de los patches de frutillas para los canales de los espacios RGB (A), HSV (B) y CIE L*a*b* (C). En todos los casos las barras de error se corresponden a la desviación estándar de las muestras (72 muestras de manchas y 100 muestras sin manchas).

$\alpha = 0,01$ indicaron que se obtienen diferencias significativas entre los valores promedio de intensidad de patches de “manchas pardas” y zonas libres manchas en todos los canales de los tres espacios estudiados. Esto implica que hay diferencias significativas en los valores de intensidad de los patches de “manchas pardas” y zonas libres de manchas ($p < 0,0001$). Si bien la representación de los datos en un espacio de color u otro puede facilitar su agrupamiento en conjuntos disjuntos más separados, el resultado del ANOVA indica que evaluando intensidad de color se pueden distinguir las “manchas pardas” de zonas no manchadas.

Los resultados de esta sección indican que las “manchas pardas” son un aspecto de calidad detectable y cuantificable, tanto en imágenes de frutillas como en el producto en sí. Este atributo de calidad está asociado a una propiedad fácilmente medible como es la intensidad media de un conjunto de píxeles en imágenes o a la percepción del color en el caso del producto fresco. Empleando imágenes de frutillas y conociendo la distribución de los valores que se corresponden a “manchas pardas”, se puede diseñar un sistema que de forma automática permita estimar la calidad de las frutillas según el grado de extensión de las “manchas pardas” en su superficie.

Capítulo 5

Conclusiones y trabajo a futuro

En este último capítulo se presenta una breve reflexión del trabajo de tesis en función de los objetivos planteados y los resultados obtenidos. En líneas generales puede decirse que el uso de técnicas de procesamiento de imágenes en alimentos es una aplicación que tiene gran potencial para el control de los procesos alimentarios. Luego de exponer las conclusiones de esta tesis, se presenta una sección denominada “Trabajo futuro”. En dicha sección, se presentan inquietudes que surgieron durante la maestría y cuyo estudio se propone continuar en los próximos años.

5.1. Conclusiones generales

En esta tesis se presentó la aplicación de técnicas de procesamiento de imágenes con el fin de extraer información visual asociada a la calidad de dos productos alimenticios. El uso de estas técnicas puede ser de gran utilidad para el control y monitoreo de alimentos, y sus procesos de transformación asociados. La información visual que se obtiene mediante el registro y procesamiento de imágenes puede vincularse con el estado del alimento, grado de avance del proceso, defectos en productos, desviaciones del proceso, entre otros aspectos. Esta información puede evitar que productos defectuosos salgan al mercado, mejorar la trazabilidad y seguridad de los productos, lograr productos más uniformes, satisfacer las expectativas del consumidor, etc. A su vez, los sistemas de obtención y procesamiento de imágenes pueden emplearse en todos los puntos de la cadena alimentaria, desde la recepción de la materia prima, su

posterior procesamiento, envasado, conservación y distribución. Sin embargo, cada aplicación en las que se puedan emplear estas técnicas requiere conocer el aspecto a determinar, el sistema de adquisición disponible y el procesamiento de información a efectuar para lograr los resultados deseados.

Si bien en esta tesis se emplearon técnicas de procesamiento que involucraron un alto grado de conocimiento del problema a resolver y operaciones particulares a las imágenes registradas, se obtuvieron resultados acordes a los objetivos planteados. En primer lugar, se incorporó la metodología empleada en el área de procesamiento de imágenes para extraer información a partir de las imágenes registradas, se aprendió la implementación y uso de diferentes algoritmos de procesamiento de imágenes junto a programas y repositorios específicos del área. También se comprendió la relevancia de operaciones transversales al procesamiento de imágenes, como lo es el registro de las mismas. Esta operación determina la calidad de la información registrada y depende de múltiples variables tales como el hardware empleado, el producto o proceso registrado, la característica a determinar, las condiciones de iluminación de la escena, la posición relativa entre la escena y el sensor que registra la imagen, entre múltiples factores. Por último, se pudo conocer la dificultad de lograr a partir de una imagen, una interpretación de una escena que sea similar a la lograda por un ser humano. Este último punto da origen a la visión artificial, una rama del conocimiento en plena expansión y que puede ser de utilidad en la automatización de las tareas de inspección en la industria alimentaria.

5.2. Conclusiones particulares

Respecto a los objetivos particulares de este trabajo, se puede concluir que la información obtenida mediante el procesamiento de imágenes se puede vincular con aspectos de calidad para cada producto en particular. Mediante un esquema tradicional de procesamiento de imágenes y empleando algoritmos disponibles en bibliotecas de código abierto, se obtuvieron resultados que permitieron evaluar las hipótesis planteadas para cada producto. Además, se planteó un procesamiento que permite una evaluación rápida de las imágenes, siendo de utilidad para un sistema automático en una línea de producción.

5.2.1. Estudio de quesos tipo *Emmental*

En el caso de quesos tipo *Emmental*, se pudo estudiar la formación de ojos a lo largo del proceso de maduración. Se presentaron dos alternativas para la estimación del volumen de ojos en quesos durante la maduración. Una de las alternativas consistió en una propuesta no invasiva basada en el uso de la geometría y masa de las hormas, junto a la densidad del queso. La segunda propuesta era una estimación semi-directa del volumen de ojos, basada en la reconstrucción espacial de dicho volumen a partir de imágenes de los quesos cortados longitudinalmente. En esta propuesta, se presentó un modelo original para realizar la reconstrucción del volumen de ojos, basado en el conocimiento de la estructura del queso. Ambas propuestas se validaron de forma conjunta al estimar de la masa del queso y contrastar este resultado con su valor experimental. A su vez, con los resultados obtenidos se modeló la evolución del volumen de ojos durante la maduración y dicho volumen se logró vincular con la altura de las hormas.

A su vez, este estudio dio lugar a dos artículos publicados en revistas científicas arbitradas del área del procesamiento de alimentos ([González et al., 2020a,b](#)) y una presentación de un trabajo completo en un congreso del área de procesamiento de señales ([González et al., 2020](#)), realizando un aporte en el área del control de alimentos. De estos tres trabajos publicados, [González et al. \(2020b\)](#) y [González et al. \(2020\)](#) están vinculados directamente al trabajo de tesis, mientras que [González et al. \(2020a\)](#) consistió en el estudio de un método acústico para la detección de la formación de ojos en quesos tipo *Emmental*.

Se considera que los resultados obtenidos en este trabajo pueden ser de utilidad en el control de calidad de quesos formadores de ojos. La propuesta presentada permite estimar de forma indirecta, rápida y no destructiva el volumen de ojos dentro de la horma al emplear la geometría y masa de las hormas junto a la densidad del queso.

5.2.2. Estudio de “manchas pardas” en frutillas

En cuanto al estudio de “manchas pardas” en frutillas, se logró determinar la capacidad de estimación visual de este atributo en frutillas frescas y en

sus imágenes por parte de un panel sensorial. Se determinó que las “manchas pardas” son un atributo que es detectado de manera consistente tanto en el producto fresco como en imágenes del mismo. Empleando la información proporcionada por el panel, se logró determinar que la intensidad media de color permite diferenciar las zonas afectadas por “manchas pardas” en imágenes de frutillas. En particular, se encontró que las “manchas pardas” se agrupan en un clúster diferente al de las zonas no manchadas en los espacios de color evaluados. Incluso, el grado de separación de ambos conjuntos depende del espacio en el que se representen los datos. A su vez, el análisis estadístico de las intensidades medias de color en los espacios evaluados indicó que existen diferencias significativas entre las zonas afectadas por “manchas pardas” y las zonas libres de este defecto. Estos resultados constituyen un primer paso en la construcción de un sistema de visión artificial para la detección automática de “manchas pardas” en frutillas.

5.3. Trabajo a futuro

Dado que esta tesis inicia una línea de investigación en el grupo “Tecnologías Aplicadas a los Procesos Alimentario” acerca de la aplicación de herramientas de procesamiento de imágenes y visión computacional en la industria alimentaria, existen múltiples ideas para implementar a futuro. A partir de las lecciones aprendidas en el marco de esta maestría, se pudo determinar la relevancia de la adquisición de las imágenes y cómo esta etapa condiciona los resultados a obtener cuando el conjunto de datos disponibles es reducido. A su vez, se pudo evaluar el potencial del uso del procesamiento de imágenes en la determinación de aspectos visuales de alimentos y como dichos aspectos se pueden vincular con su estado.

En las actividades a realizar, se propone continuar con esta línea de investigación, efectuando un estudio más detallado de los parámetros que afectan la adquisición de las imágenes y que acentúan las características cuya información se quiere extraer. En particular, se propone estudiar el efecto de la iluminación a la hora de obtener las imágenes, variando el tipo de luz empleado y la posición relativa entre el objeto y la cámara. Otro aspecto que se desea profundizar es la determinación de color y textura de alimentos a partir de

fotografías, contrastando estas medidas con la percepción visual del consumidor, su medición instrumental y la evolución de dichas propiedades durante el procesamiento de los alimentos. En el caso de los alimentos estudiados, se planea realizar una validación de los modelos presentados para queso tipo *Emmental* a nivel industrial. Mientras que para el caso de frutillas se planifica la publicación de un artículo científico con los resultados obtenidos, realizando así un aporte de acerca de la determinación de aspectos de calidad en frutillas mediante procesamiento de imágenes.

Referencias bibliográficas

- Aharchi, M. and Kbir, M. A. (2019). A review on 3d reconstruction techniques from 2d images. In The Proceedings of the Third International Conference on Smart City Applications, pages 510–522. Springer.
- Ambrosio, L. and Tortorelli, V. M. (1990). Approximation of functional depending on jumps by elliptic functional via t-convergence. Communications on Pure and Applied Mathematics, 43(8):999–1036.
- Ares, G., Barrios, S., Lareo, C., and Lema, P. (2009). Development of a sensory quality index for strawberries based on correlation between sensory data and consumer perception. Postharvest biology and technology, 52(1):97–102.
- Ballard, D. H. (1981). Generalizing the hough transform to detect arbitrary shapes. Pattern recognition, 13(2):111–122.
- Barten, H. M. and Pfister, F. K. (1980). Method for the contactless determination of features of meat quality. US Patent 4,226,540.
- Bhat, R., Geppert, J., Funken, E., and Stamminger, R. (2015). Consumers perceptions and preference for strawberries—a case study from germany. International Journal of Fruit Science, 15(4):405–424.
- Bradski, G. and Kaehler, A. (2000). Opencv. Dr. Dobb’s journal of software tools, 3.
- Brosnan, T. and Sun, D.-W. (2004). Improving quality inspection of food products by computer vision—a review. Journal of Food Engineering, 61(1):3 – 16. Applications of computer vision in the food industry.
- Caccamo, M., Melilli, C., Barbano, D., Portelli, G., Marino, G., and Licitra, G. (2004). Measurement of gas holes and mechanical openness in cheese by image analysis. Journal of dairy science, 87(3):739–748.

- Cangar, ., Leroy, T., Guarino, M., Vranken, E., Fallon, R., Lenehan, J., Mee, J., and Berckmans, D. (2008). Automatic real-time monitoring of locomotion and posture behaviour of pregnant cows prior to calving using online image analysis. Computers and Electronics in Agriculture, 64(1):53 – 60.
- Canny, J. (1986). A computational approach to edge detection. IEEE Transactions on pattern analysis and machine intelligence, 8(6):679–698.
- Castro-Fontoura, G. (2016). Productividad y agrotecnología en uruguay: Desafíos para el crecimiento de la productividad en el sector agropecuario y posibles soluciones.
- Chen, Y.-R., Chao, K., and Kim, M. S. (2002). Machine vision technology for agricultural applications. Computers and Electronics in Agriculture, 36(2):173 – 191.
- Cho, W., Na, M., Kim, S., and Jeon, W. (2019). Automatic prediction of brix and acidity in stages of ripeness of strawberries using image processing techniques. In 2019 34th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), pages 1–4. IEEE.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. psychometrika, 16(3):297–334.
- Cubero, S., Aleixos, N., Moltó, E., Gómez-Sanchis, J., and Blasco, J. (2011). Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables. Food and bioprocess technology, 4(4):487–504.
- Davies, E. R. (2000). Image processing for the food industry, volume 37. World Scientific.
- Du, C.-J. and Sun, D.-W. (2004). Recent developments in the applications of image processing techniques for food quality evaluation. Trends in food science & technology, 15(5):230–249.
- Duggal, V., Sukhwani, M., Bipin, K., Reddy, G. S., and Krishna, K. M. (2016). Plantation monitoring and yield estimation using autonomous quadcopter for precision agriculture. In 2016 IEEE international conference on robotics and automation (ICRA), pages 5121–5127. IEEE.

- ElMsry, G., ElSheikh, I., and Morsy, N. (2009). Colour grading of strawberry using computer vision and backpropagation artificial neural network. Mansoura University Journal of Agricultural Sciences.
- Eskelinen, J., Alavuotunki, A., Hæggström, E., and Alatossava, T. (2007). Preliminary study of ultrasonic structural quality control of swiss-type cheese. Journal of dairy science, 90(9):4071–4077.
- FINNEY JR, E. E. (1978). Engineering techniques for nondestructive quality evaluation of agricultural products1. Journal of food protection, 41(1):57–62.
- Fox, P. F., Guinee, T. P., Cogan, T. M., and McSweeney, P. L. (2017). Fundamentals of cheese science. Springer.
- Fröhlich-Wyder, M. and Bachmann, H. (2007). Swiss cheese. In McSweeney, P. L. H., editor, Cheese problems solved. Woodhead Publishing Limited.
- Golnabi, H. and Asadpour, A. (2007). Design and application of industrial machine vision systems. Robotics and Computer-Integrated Manufacturing, 23(6):630–637.
- González, M., Budelli, E., Pérez, N., and Lema, P. (2020a). Acoustic techniques to detect eye formation during ripening of emmental type cheese. Innovative Food Science & Emerging Technologies, 59:102270.
- González, M., Budelli, E., Pérez, N., and Lema, P. (2020b). Estimation of eye formation in swiss-type cheese using geometrical measurements. Journal of Food Engineering, 275:109848.
- Gonzalez, R. C. and Woods, R. E. (2002). Digital Image Processing. Prentice-Hall Inc.
- González, M., Budelli, E., Pérez, N., and Lema, P. (2020). Image processing applied to eye segmentation in cheese maturation. In 2020 IEEE International Symposium on Circuits and Systems (ISCAS), pages 1–3.
- Gunasekaran, S. (1996). Computer vision technology for food quality assurance. Trends in Food Science & Technology, 7(8):245–256.
- Han, K.-S., Kim, S.-C., Lee, Y.-B., Kim, S.-C., Im, D.-H., Choi, H.-K., and Hwang, H. (2012). Strawberry harvesting robot for bench-type cultivation. Journal of Biosystems Engineering, 37(1):65–74.

- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., et al. (2020). Array programming with numpy. Nature, 585(7825):357–362.
- He, L., Ren, X., Gao, Q., Zhao, X., Yao, B., and Chao, Y. (2017). The connected-component labeling problem: A review of state-of-the-art algorithms. Pattern Recognition, 70:25–43.
- Hélias, A., Mirade, P.-S., and Corrieu, G. (2007). Modeling of camembert-type cheese mass loss in a ripening chamber: Main biological and physical phenomena. Journal of Dairy Science, 90(11):5324–5333.
- Huc, D., Challoy, S., Monziols, M., Michon, C., and Mariette, F. (2014a). Spatial characterisation of eye-growing kinetics in semi-hard cheeses with propionic acid fermentation. International Dairy Journal, 39(2):259–269.
- Huc, D., Mariette, F., Challoy, S., Barreau, J., Moulin, G., and Michon, C. (2014b). Multi-scale investigation of eyes in semi-hard cheese. Innovative Food Science & Emerging Technologies, 24:106–112.
- Huc, D., Moulin, G., Mariette, F., and Michon, C. (2013). Investigation of curd grains in swiss-type cheese using light and confocal laser scanning microscopy. International Dairy Journal, 33(1):10–15.
- Hutchings, J. B. (1977). The importance of visual appearance of foods to the food processor and the consumer 1. Journal of Food Quality, 1(3):267–278.
- Indrabayu, I., Arifin, N., and Areni, I. S. (2019). Strawberry ripeness classification system based on skin tone color using multi-class support vector machine. In 2019 International Conference on Information and Communications Technology (ICOIACT), pages 191–195. IEEE.
- Knudsen, E. S. (1970). Methods of determining the quality of meat and devices for carrying out said methods. US Patent 3,493,774.
- Kondo, N. (2010). Automation on fruit and vegetable grading system and food traceability. Trends in Food Science & Technology, 21(3):145–152.
- Kraggerud, H., Wold, J., Høy, M., and Abrahamsen, R. (2009). X-ray images for the control of eye formation in cheese. International journal of dairy technology, 62(2):147–153.

- Law, B. A. and Tamime, A. Y. (2011). Technology of cheesemaking, volume 18. John Wiley & Sons.
- LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11):2278–2324.
- Lewers, K. S., Newell, M. J., Park, E., and Luo, Y. (2020). Consumer preference and physiochemical analyses of fresh strawberries from ten cultivars. International Journal of Fruit Science, pages 1–24.
- Li, J., Tan, J., and Shatadal, P. (2001). Classification of tough and tender beef by image texture analysis. Meat science, 57(4):341–346.
- Liming, X. and Yanchao, Z. (2010). Automated strawberry grading system based on image processing. Computers and Electronics in Agriculture, 71:S32–S39.
- Liu, Q., Sun, K., Peng, J., Xing, M., Pan, L., and Tu, K. (2018). Identification of bruise and fungi contamination in strawberries using hyperspectral imaging technology and multivariate analysis. Food analytical methods, 11(5):1518–1527.
- Lukinac, J., Jukić, M., Mastanjević, K., and Lučan, M. (2018). Application of computer vision and image analysis method in cheese-quality evaluation: a review. Ukrainian food journal, 2(7):192–214.
- Marr, D. and Poggio, T. (1979). A computational theory of human stereo vision. Proceedings of the Royal Society of London. Series B. Biological Sciences, 204(1156):301–328.
- Marr, D. and Ullman, S. (1981). Directional selectivity and its use in early visual processing. Proceedings of the Royal Society of London. Series B. Biological Sciences, 211(1183):151–180.
- Matar, C., Gaucel, S., Gontard, N., Guilbert, S., and Guillard, V. (2018). A global visual method for measuring the deterioration of strawberries in map. MethodsX, 5:944–949.
- McAndrew, A. (2015). A computational introduction to digital image processing. CRC Press.

- Mumford, D. B. and Shah, J. (1989). Optimal approximations by piecewise smooth functions and associated variational problems. Communications on pure and applied mathematics.
- Musse, M., Challos, S., Huc, D., Quellec, S., and Mariette, F. (2014). Mri method for investigation of eye growth in semi-hard cheese. Journal of Food Engineering, 121:152–158.
- Næs, T., Brockhoff, P. B., and Tomić, O. (2010). Statistics for sensory and consumer science. Wiley Online Library.
- Nagata, M., Tallada, J. G., and Kobayashi, T. (2006). Bruise detection using nir hyperspectral imaging for strawberry (*fragaria* × *ananassa* duch.). Environmental Control in Biology, 44(2):133–142.
- Nunes, J. L., Piquerez, M., Pujadas, L., Armstrong, E., Fernández, A., and Lecumberry, F. (2015). Beef quality parameters estimation using ultrasound and color images. BMC bioinformatics, 16(S4):S6.
- Nunnally, J. C. (1967). Psychometric Theory. McGraw-Hill.
- Observatorio Granjero (2017). Resultados de evaluación de pérdidas en la cadena agroalimentaria de frutilla.
- Oficina Estadísticas Agropecuarias (2019). Estadísticas del sector lácteo 2018.
- Oo, L. M. and Aung, N. Z. (2018). A simple and efficient method for automatic strawberry shape and size estimation and classification. Biosystems engineering, 170:96–107.
- Otsu, N. (1979). A threshold selection method from gray-level histograms. IEEE transactions on systems, man, and cybernetics, 9(1):62–66.
- Patel, K. K., Kar, A., Jha, S., and Khan, M. (2012). Machine vision system: a tool for quality inspection of food and agricultural products. Journal of food science and technology, 49(2):123–141.
- Pathare, P. B., Opara, U. L., and Al-Said, F. A.-J. (2013). Colour measurement and analysis in fresh and processed foods: a review. Food and bioprocess technology, 6(1):36–60.

- Pinder, A. C. and Godfrey, G. (2012). Food process monitoring systems. Springer Science & Business Media.
- Ramirez, M. M. G., Rincon, J. C. V., and Parada, J. F. L. (2014). Liquid level control of coca-cola bottles using an automated system. In 2014 International Conference on Electronics, Communications and Computers (CONIELECOMP), pages 148–154. IEEE.
- Ribeiro, F. D. S., Gong, L., Calivá, F., Swainson, M., Gudmundsson, K., Yu, M., Leontidis, G., Ye, X., and Kollias, S. (2018). An end-to-end deep neural architecture for optical character verification and recognition in retail food packaging. In 2018 25th IEEE International Conference on Image Processing (ICIP), pages 2376–2380. IEEE.
- Ridler, T. and Calvard, S. (1978). Picture thresholding using an iterative selection method. IEEE transactions on systems, man, and cybernetics, 8(8):630–632.
- Sampson, D. J., Chang, Y. K., Rupasinghe, H. V., and Zaman, Q. U. (2014). A dual-view computer-vision system for volume and image texture analysis in multiple apple slices drying. Journal of Food Engineering, 127:49–57.
- Savant, S. (2014). A review on edge detection techniques for image segmentation. International Journal of Computer Science and Information Technologies, 5(4):5898–5900.
- Schuetz, P., Guggisberg, D., Jerjen, I., Fröhlich-Wyder, M., Hofmann, J., Wechsler, D., Flisch, A., Bisig, W., Sennhauser, U., and Bachmann, H.-P. (2013). Quantitative comparison of the eye formation in cheese using radiography and computed tomography data. International Dairy Journal, 31(2):150–155.
- Serra, J. and Soille, P. (2012). Mathematical morphology and its applications to image processing, volume 2. Springer Science & Business Media.
- Sezgin, M. and Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. Journal of Electronic imaging, 13(1):146–166.

- Simal, S., Sánchez, E., Bon, J., Femenia, A., and Rossello, C. (2001). Water and salt diffusion during cheese ripening: effect of the external and internal resistances to mass transfer. Journal of food Engineering, 48(3):269–275.
- Simpson, J. E. (1965). Quality evaluation studies of foreign and domestic rices. US Department of Agriculture.
- Sun, D.-W. (2016). Computer vision technology for food quality evaluation. Academic Press.
- Suzuki, S. et al. (1985). Topological structural analysis of digitized binary images by border following. Computer vision, graphics, and image processing, 30(1):32–46.
- Teubner, U. and Brückner, H. J. (2019). Optical Imaging and Photography: Introduction to Science and Technology of Optics, Sensors and Systems. Walter de Gruyter GmbH & Co KG.
- Tillett, R. (1991). Image analysis for agricultural processes: a review of potential opportunities. Journal of agricultural Engineering research, 50:247–258.
- Tönshoff, H., Janocha, H., and Seidel, M. (1988). Image processing in a production environment. CIRP annals, 37(2):579–590.
- Trienekens, J. and Zuurbier, P. (2008). Quality and safety standards in the food industry, developments and challenges. International Journal of Production Economics, 113(1):107–122.
- Trussell, H. J. and Vrhel, M. J. (2008). Fundamentals of digital imaging. Cambridge University Press.
- Uruguay XXI (2020). Agronegocios.
- US Department Agriculture (2001). Swiss Cheese, Emmentaler Cheese Grades and Standards.
- Van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., and Yu, T. (2014). scikit-image: image processing in python. PeerJ, 2:e453.
- Van Rossum, G. and Drake, F. L. (2011). The python language reference manual. Network Theory Ltd.

- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., et al. (2020). Scipy 1.0: fundamental algorithms for scientific computing in python. Nature methods, 17(3):261–272.
- Wadhvani, R. and McMahon, D. (2012). Color of low-fat cheese influences flavor perception and consumer liking. Journal of dairy science, 95(5):2336–2346.
- Wrangborg, D. and Ardö, H. (2007). Automatic fault detection in cheese using computer vision. Lund University.
- Wyszecki, G. and Stiles, W. S. (1982). Color science, volume 8. Wiley New York.
- Yu, Y., Zhang, K., Yang, L., and Zhang, D. (2019). Fruit detection for strawberry harvesting robot in non-structural environment based on mask-rcnn. Computers and Electronics in Agriculture, 163:104846.
- Zareiforush, H., Minaei, S., Alizadeh, M. R., and Banakar, A. (2015). Potential applications of computer vision in quality inspection of rice: a review. Food engineering reviews, 7(3):321–345.
- Zhang, B., Huang, W., Li, J., Zhao, C., Fan, S., Wu, J., and Liu, C. (2014). Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: A review. Food Research International, 62:326–343.