

Proyecto Fin de Carrera

Detección automática de objetos extraños
utilizando Rayos X basados en interferometría
por rejillas de difracción

Autor

Mónica Emerson Serrano

Director: Line Katrine Harder Clemmensen

Ponente: José Elías Herrero Jaraba

Escuela de Ingeniería y Arquitectura.

Junio 2014

Detección automática de objetos extraños utilizando Rayos X basados en interferometría por rejillas de difracción. RESUMEN

Los controles de calidad en la industria alimenticia son esenciales tanto por seguridad como de cara a la satisfacción del cliente. Durante el proceso de fabricación de comida, objetos no deseados pueden introducirse en los productos, resultando peligroso o poco apetitoso para el consumidor. Hoy en día, los sistemas de rayos-X incorporados a líneas de producción pueden detectar materia no orgánica. Sin embargo, encontrar objetos orgánicos en comida utilizando rayos-X convencionales todavía supone un reto.

El objetivo de este proyecto es demostrar la mejora introducida por una nueva técnica de rayos-X en la detección de objetos extraños, cuando éstos pueden ser de origen orgánico. Esta novedosa técnica está basada en interferometría de rayos, creada a base de añadir rejillas de difracción a una fuente de rayos-X convencional. Así, obtenemos información sobre las propiedades de absorción, refracción y dispersión de una muestra; mientras que los rayos-X convencionales sólo proporcionan la absorción.

Los datos fueron tomados personalmente en *Technische Universität München*. Cada imagen consta de tres modalidades (absorción, contraste de fase y campo oscuro) y contiene una muestra de comida contaminada por objetos de distintos tamaños, orgánicos y no orgánicos. Se tomaron imágenes de varios productos que cuentan con distintas propiedades y son de importancia para los colaboradores del proyecto **NEXIM** (New X-ray Imaging Modalities for safe and high quality food) dentro del cual se encuentra este PFC.

En esta memoria, se compara el rendimiento de dos métodos de clasificación, uno supervisado y otro no-supervisado. Se hará hincapié en la técnica no-supervisada, comparando modelos de comida con distintos atributos y contrastando los resultados de detección con aquellos que se obtendrían usando únicamente la modalidad de absorción.

Palabras clave

Rayos-X, refracción, dispersión, detección, GMM, SVM, estadística, imagen, aprendizaje de máquinas, objetos extraños.

Prefacio

Este proyecto fin de carrera ha sido desarrollado en el departamento de Matemáticas Aplicadas e Informática de la Universidad Técnica de Dinamarca, *Danmarks Tekniske Universitet* (DTU).

El proyecto ha sido supervisado por Line Clemmensen y Hildur Einarsdottir, ambas del departamento DTU Compute, y desarrollado entre el 20 de Enero de 2014 y el 12 de Junio de 2014 a tiempo completo.

El proyecto trata la detección de objetos extraños en comida, utilizando una novedosa técnica de rayos-X.

Lyngby, 12-Junio-2014

A handwritten signature in black ink. It features a large, stylized 'S' shape that loops around the word 'Emerson' and extends to the right, ending with a flourish. The name 'Mónica' is written in a smaller, cursive script to the right of the main signature.

Monica J. Emerson

Agradecimientos

Me gustaría dar las gracias a mis dos supervisoras, Line y Hildur. Sin Line no hubiera tenido la oportunidad de hacer este proyecto y Hildur ha sido una parte esencial durante el desarrollo del proyecto; compartiendo sus conocimientos y haciendo posible la toma de datos directa y personal en *Technische Universität München* (TUM). Además, los colaboradores del proyecto **NEXIM** merecen reconocimiento por financiar mi viaje a Munich.

Los miembros del grupo de análisis de imagen de DTU Compute también se merecen mi agradecimiento, por los consejos tan valiosos que me dieron mientras realizaba mi proyecto y, en general, dar las gracias también a todos los miembros del Área de análisis de datos y estadística, por hacerme sentir bienvenida. Especialmente a Ruta, que me prestó su mesa, Merete, que apreció mi trabajo y David, que sabe sacarte una sonrisa hasta en el peor de tus momentos.

También estoy muy agradecida a la gente del departamento de Física de TUM. Kai tuvo que reservar el montaje experimental para que yo lo pudiese usar y Konstantin estaba atento para ayudar siempre que fuese necesario.

Finalmente, el proyecto no hubiese sido posible sin el apoyo financiero de mis padres. Mi padre Rupert fue el corrector principal de la versión inglesa del proyecto. Además, siempre está ahí para escuchar y propocionar consejos de cara a mi carrera profesional, demostrando sus habilidades como consultor. Gracias a mi hermana Rebeca, a mi madre María Ángeles y a mi amigo José Antonio por leer y corregir la versión final. Y, especialmente, gracias a mi madre por dedicar su vida a sus hijas.

Abreviaturas

BIF	Atributos básicos de imagen
CM	Matriz de Confusión
EM	Algoritmo Esperanza-Maximización
FS	Selección de atributos
FP	Falsos Positivos
FPR	Razón de Falsos Positivos o razón de falsa alarma
GBI	Interferometría basada en rejillas de difracción
GPU	Unidad de Procesamiento Gráfico
GMM	Modelo mezcla de Gaussianas
ROC	Característica Operativa del Receptor
SVM	Máquina de Vectores de Soporte
TP	Verdaderos positivos
TPR	Razón de verdaderos positivos
TUM	<i>Technische Universität München</i>

Tabla de contenido

Prefacio	i
Agradecimientos	iii
Abreviaturas	v
1 Introducción	1
1.1 Motivación	1
1.2 Problemática	2
1.3 Estudio de la bibliografía	5
1.3.1 ¿Cómo encuentran los rayos-X objetos extraños?	6
1.3.1.1 ¿Qué son los rayos-X?	6
1.3.1.2 Distintas estrategias para encontrar objetos extraños	7
1.4 Organización de la memoria	8
2 Diseño	9
2.1 Montaje experimental	9
2.2 Materiales	12
2.2.1 Comida	13
2.2.2 Objetos extraños	14
2.3 Métodos	17
2.3.1 Preprocesado	17
2.3.2 Atributos	18
2.3.2.1 Atributos de textura	18
2.3.2.2 Regresión/Selección de atributos (FS)	19
2.3.3 Métodos de entrenamiento y clasificación	20
2.3.3.1 GMM	20
2.3.3.2 SVM	22

3	Resultados	25
3.1	Conjunto de datos principal: Pavo	26
3.1.1	Modelos Gaussianos	26
3.1.2	Modelo Gaussiano VS Máquina de Vectores de Soporte .	37
3.2	Conjunto de datos ampliado: Queso, Salami y Pan de centeno. .	39
4	Conclusiones y trabajo futuro	49
	Referencias	53

CAPÍTULO 1

Introducción

El primer capítulo explica la necesidad de detectar objetos extraños y, posteriormente, introduce el reto que supone detectar objetos compuestos por materia orgánica. Después de hablar de los antecedentes de esta nueva técnica de escaneado por rayos-X, se tratará el "state of the art" en detección de objetos extraños, es decir, cuáles son las últimas técnicas que han sido investigadas para detectar objetos extraños.

1.1 Motivación

La detección de objetos extraños es de gran importancia en los procesos de control de calidad de comida. Nunca es agradable ni higiénico encontrarse una araña o una piedra en la comida. Hacerte daño mientras comes podría crear muy mala propaganda para el productor de ese alimento en cuestión, lo cual incluso podría acarrear serios problemas judiciales. Por ello, el control de calidad es una tarea de alta prioridad en la industria alimenticia.

Este proyecto proporcionará respuesta a dos preguntas principales:

1. ¿Cómo podemos inspeccionar los productos alimenticios de una manera automática y eficiente?
2. ¿Podemos encontrar materiales orgánicos - como insectos, plástico o madera - utilizando rayos-X?

1.2 Problemática

En numerosos campos existe la necesidad de obtener información sobre la estructura interna de objetos complejos sin haber sido diseccionados. Gracias a los rayos-X, podemos obtener esta información de forma no invasiva. Esta información estructural es útil en una amplia gama de industrias para inspección, ya sea para control de calidad de productos, para diagnóstico médico o para controles de seguridad, entre otras aplicaciones.

Los sistemas de rayos-X convencionales han sido usados durante años para la detección en línea de objetos extraños en comida. El problema es que los rayos-X por sí solos no sirven para detectar diferencias entre materiales blandos. Por tanto, la detección de objetos orgánicos en comida supone un reto.

Gracias a nuevos métodos de escaneo con rayos-X, podemos medir, no sólo la absorción de una muestra, sino también sus propiedades de refracción y dispersión. Así, aumentamos el contraste entre objetos blandos, pudiendo diferenciar comida de objetos extraños por el mero hecho de contar con distintas propiedades de refracción y/o dispersión. Como consecuencia, aumenta enormemente el número de escenarios en los que la información estructural de los objetos puede ser obtenida, incrementando el número de aplicaciones posibles.

El objetivo de este estudio es detectar objetos extraños - objetos que no deberían estar en un lugar específico (concretamente en comida) - utilizando una nueva técnica de rayos-X conocida como interferometría basada en rejillas de difracción (GBI). Usando esta nueva técnica, se obtienen imágenes multi-modales de comida contaminada con objetos extraños. Las tres modalidades de imagen son transmisión, contraste de fase y campo oscuro, que se corresponden con las propiedades de absorción, refracción y dispersión, respectivamente.

Para investigar la aplicabilidad de esta técnica en la detección de objetos extraños en comida, varios productos alimenticios son elegidos junto con una se-

lección de objetos extraños que son encontrados típicamente en la industria alimentaria.

La Figura 1.1 es un resumen de las sugerencias proporcionadas por colaboradores daneses sobre qué objetos contaminantes suelen encontrar y de qué dimensiones en cada tipo de industria alimenticia.

Food Product	Industrial Partner	Contaminant	Size distribution
Minced meat	Danish Crown	Bone Plastic Cartilage	0.5 mm, 1 mm, 5 mm, 10 mm, 50 mm, 100 mm, 1 cm, 5 cm, 10 cm
Bread	Lantmannen Schulstad Kohberg	Stones	?
Milk products	Arla	?	?
Chicken products	Lantmännen Danpo	Bone Cartilage Plastic	0.5 mm, 1 mm, 5 mm, 10 mm, 50 mm, 100 mm, 1 cm, 5 cm, 10 cm
Pet Food	Arovit Petfood	?	?
Spring rolls	Daloon	String Stones	?

Figura 1.1: Sugerencias sobre productos alimenticios, objetos extraños y su tamaño, proporcionadas por colaboradores industriales.

En la Figure 1.2 podemos observar, dentro del pan de centeno, los distintos materiales seleccionados como objetos extraños.



Figura 1.2: Verde: Madera, Azul: Vidrio, Rosa: Plástico duro, Morado: Plástico blando, Amarillo: Goma, Turquesa: Piedra, Naranja: Metal y Rojo: Insecto.

Esta novedosa técnica podría permitir que los objetos orgánicos destaquen frente a la comida, también orgánica, gracias a la diferencia entre las propiedades de refracción y/o dispersión de comida y objetos. Esto no era posible sólo con la absorción, ya que los objetos orgánicos, al igual que la comida, tienen agua como componente principal. Por ello, presentan una atenuación muy similar y, consiguientemente, no existe diferencia de contraste entre comida y objetos extraños en la modalidad típica de los rayos-X (absorción).

En la Figura 1.3 podemos ver que distintos materiales destacan mejor en distintas modalidades. El vidrio, el metal y las piedras, pueden ser detectadas fácilmente por la diferencia en intensidad de absorción con respecto a la comida. Por el contrario, el plástico (duro y blando), la goma, los insectos y la madera, destacan más en las otras dos modalidades.

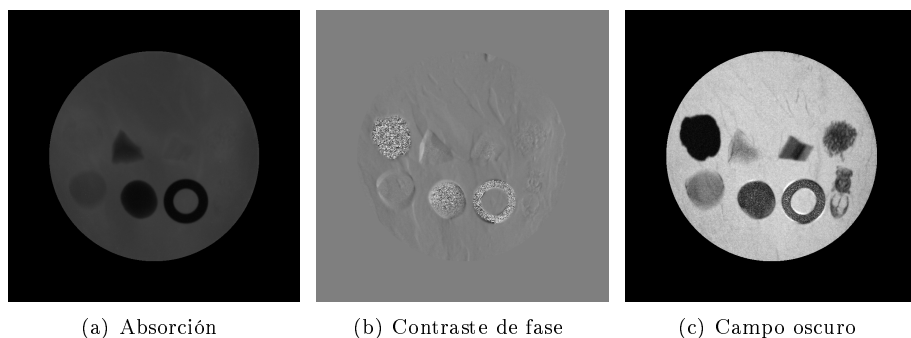


Figura 1.3: Filete de ternera con objetos extraños para las tres modalidades. Los objetos extraños se ven con distinto contraste dependiendo de la modalidad.

Resumiendo, para detectar objetos extraños en comida, se montará un caso de prueba. Para la realización de este caso de prueba, son necesarios los siguientes pasos.

1. Los datos tienen que ser recogidos, por lo que los productos alimenticios y los objetos a investigar deben ser definidos, además de limitados para que la toma de datos sea razonable.
2. Para analizar los datos adquiridos, dos algoritmos distintos serán investigados y sus parámetros serán optimizados.
3. Se comparará la valía de estos algoritmos a través de los resultados obtenidos.

4. Se mostrará la mejora en eficiencia introducida por esta nueva técnica de rayos-X en la detección de objetos extraños en comida.

Antes de entrar en detalles del caso de prueba específico, se tratarán los antecedentes y las técnicas más recientes en cuanto a la detección de objetos.

1.3 Estudio de la bibliografía

Las técnicas de escaneo han sido esenciales en numerosos campos. De hecho, cada vez que se inventaba una nueva técnica, solía ir seguida de un desarrollo científico. Después del descubrimiento de los rayos-X por Wilhelm Röntgen en 1895 [Röntgen, 1895], los rayos-X se utilizaron para medicina [Spiegel, 1995]. Similarmente, tras la invención del primer microscopio óptico en el siglo XVII, se investigaron muestras biológicas [Hooke, 1665].

Durante el siglo XX, aparecieron nuevas formas de escaneo, como el contraste de fase y el campo oscuro [Rost and Oldfield, 2000, Zernike, 1942]. Estas nuevas modalidades miden la dispersión y la refracción, respectivamente, en lugar de basarse en la absorción.

En la última década, ha habido una evolución enorme en toma de imágenes con rayos-X, apareciendo estas dos nuevas modalidades de imagen, que son más sensibles que la absorción a materiales de baja impedancia, como la materia orgánica. La interferometría por rejillas de difracción permite obtener estas tres modalidades (absorción, contraste de fase y campo oscuro) a la vez y con correspondencia entre píxeles. De momento, puede ser implementada en laboratorios [Pfeiffer et al., 2006], pero parece prometedora de cara a ser implementada industrialmente [Kottler et al., 2010].

Varios estudios han sido llevados a cabo sobre esta reciente técnica de imagen, sus parámetros y aplicaciones [Nielsen, 2012, Bech, 2009]. Además de esta nueva técnica, se han considerado otras estrategias para aumentar el contraste en absorción entre objetos blandos, como la utilización de energías en el rango de 10 a 25 keV para aprovechar la mayor atenuación que presentan los elementos ligeros en este espectro [Nielsen et al., 2012]. Finalmente, se decidió que merecía la pena investigar la técnica basada en interferometría por rejillas de difracción, por no implicar una velocidad lenta de escaneo ni potencias tan altas de rayos-X.

1.3.1 ¿Cómo encuentran los rayos-X objetos extraños?

1.3.1.1 ¿Qué son los rayos-X?

Los rayos-X son una forma de radiación electromagnética invisible, de la misma naturaleza que las ondas de radio, las microondas, la luz infrarroja y visible, y los rayos Gamma y ultravioleta. La diferencia principal entre los rayos-X y los rayos Gamma reside en su origen: los primeros son producidos por eventos extra-nucleares, mientras que los segundos provienen del núcleo. La energía de los rayos-X está entre la de los rayos ultravioleta y los rayos Gamma, tal y como podemos ver en la Figura 1.4.

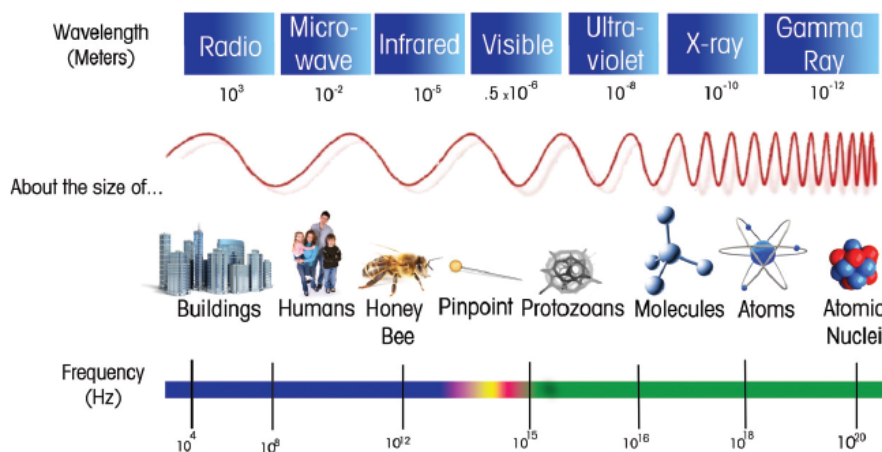


Figura 1.4: Espectro electromagnético [Toledo, 2014].

El hecho de que los rayos-X no pasen con la misma facilidad por todos los materiales, es lo que produce la diferencia en intensidad observada en la modalidad de transmisión. Esta habilidad para penetrar los materiales, también depende de la densidad y el grosor del material en cuestión. Tal y como se puede ver en la Figura 1.5, cuanto más espeso es el queso, es decir, cuantas más lonchas hay apiladas, más difícil es traspasarlo. Como consecuencia, se atenúan más los rayos y menor cantidad de señal llega al detector, resultando en una intensidad más oscura que la del queso. Por consiguiente, los objetos extraños que son más densos que el queso destacan en la modalidad de absorción.

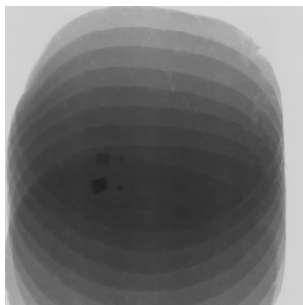


Figura 1.5: Absorción de lonchas de queso apiladas, conteniendo varios objetos extraños.

1.3.1.2 Distintas estrategias para encontrar objetos extraños

La visión por computador y la visión humana son muy distintas, a pesar de que ambas se basan en imágenes. Por ello es difícil, pero necesario, evitar que nuestras ideas preconcebidas influyan el tipo de procesamiento de imagen que debe ser aplicado cuando buscamos objetos extraños.

Algunos estudios han intentado encontrar cuerpos extraños en imágenes de absorción [Batchelor et al., 2004]. Estas técnicas se apoyan en las diferencias de intensidad en la imagen, creadas por la diferencia de densidad entre los objetos extraños y el producto alimenticio. Estas intensidades pueden ser homogéneas para todo el producto u objeto en cuestión, o pueden seguir un cierto patrón, como la carne picada de la Figura 1.6. Por esta razón, también son importantes los estudios que abarcan el análisis de texturas. Más aún, también se podrían analizar las formas de los objetos, pero en este caso de prueba concreto no es de utilidad, ya que los objetos fragmentados podrían tener infinitas formas distintas.

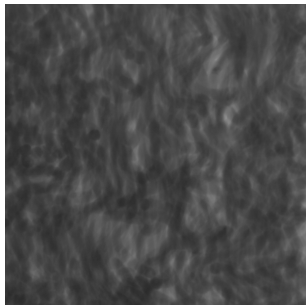


Figura 1.6: Absorción de carne picada sin objetos extraños.

Este proyecto se centra en combinar la nueva técnica de interferometría de rayos-X (GBI) con análisis de datos basado en técnicas estadísticas, de análisis de imagen y de aprendizaje de máquinas; incluyendo información tanto espectral (intensidad) como espacial (textura).

1.4 Organización de la memoria

En líneas generales, el resto de la memoria trata los siguientes aspectos:

1. El capítulo 2 se centra en el aspecto de diseño. Primero, se describe rápidamente el montaje experimental de rayos-X utilizado para la toma de datos. Segundo, se argumenta la elección de los materiales (comida y objetos extraños). Y, tercero, se explican los métodos aplicados y se analizan las razones que han llevado a su selección.
2. En el tercer capítulo, se presentan los resultados obtenidos. En primer lugar, se comparan distintos modelos Gaussianos para un producto concreto. En segundo lugar, se evalúa el rendimiento del modelo Gaussiano frente al de las máquinas de vectores de soporte (SVM) para un conjunto de atributos concreto. Por último, se investigan varios productos alimenticios.
3. En el último capítulo, se extraen conclusiones relativas a los resultados y se trata el trabajo futuro.

CAPÍTULO 2

Diseño

El siguiente capítulo examina la faceta de diseño de la tarea de detección, tratando la técnica de escaneado basada en rejillas de difracción y los métodos de detección.

2.1 Montaje experimental

En los rayos-X convencionales, para calcular las propiedades de absorción de una muestra se mide el cambio de amplitud en el haz de rayos, lo cual está representado en la Figura 2.1, donde el perfil del haz verde muestra una disminución en amplitud tras atravesar un material absorbente. Cuando se amplía el montaje, añadiendo las rejillas de difracción, se crea un patrón de interferencias. Analizando este patrón de interferencias, se obtiene la refracción y dispersión de la muestra, además de la absorción. Estas propiedades de refracción y dispersión se corresponden en el detector con un desplazamiento o ensanchamiento del haz, respectivamente. Esto es ilustrado en la Figura 2.1 por los perfiles de haz azul y rojo.

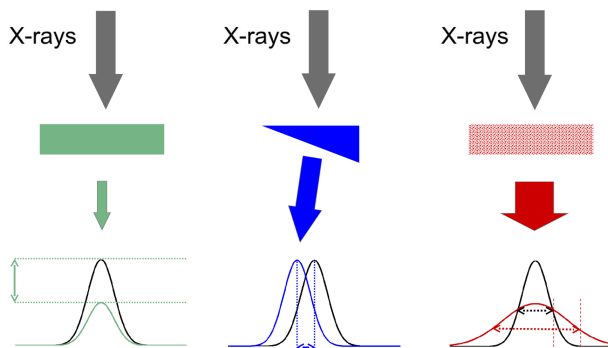


Figura 2.1: Interacciones: Cambio que sufre un haz incidente, tras insertar un material en color *Verde*: absorbente *Azul*: refractivo *Rojo*: posee una microestructura ordenada [Nielsen, 2012].

En la Figura 2.2 se muestra el montaje experimental del GBI. La rejilla G1 produce un patrón de interferencias periódico y transversal a la dirección de los rayos. El cambio en amplitud o posición de este patrón es investigado con una segunda rejilla G2 parte por parte (moviendo una de las rejillas, G1 en este montaje específico), obteniendo medidas simultáneas de las propiedades de absorción, refracción y dispersión de la muestra a estudiar. La tercera rejilla G0 añade coherencia espacial al haz de rayos-X [Nielsen, 2012, Bech, 2009].

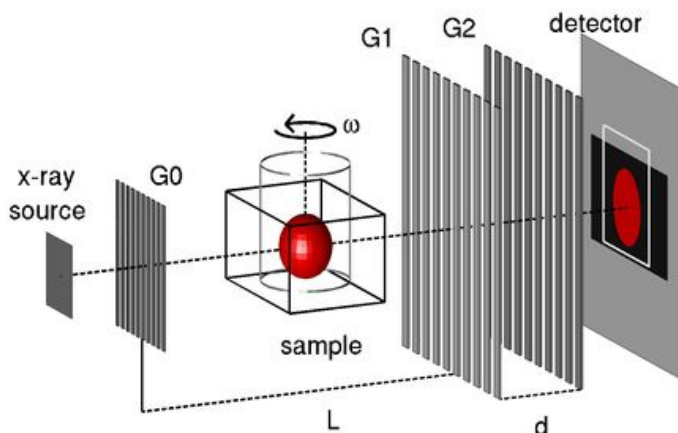


Figura 2.2: Esquema de un interferómetro Talbot-Lau [Bech, 2009]

El sistema de rayos-X basado en rejillas de difracción (GBI) que ha sido utilizado para adquirir las imágenes de las muestras de comida, está situado en el departamento de Física Biomédica de *Technische Universität München* (TUM). Es un interferómetro Talbot-Lau, el cual combina un tubo de rayos-X convencional y tres rejillas de difracción (fuente, de fase y analizadora), fabricadas por Microworks (Karlsruhe, Germany) con periodos entre rendijas de $10\text{ }\mu\text{m}$, $3.24\text{ }\mu\text{m}$ y $4.8\text{ }\mu\text{m}$ respectivamente. La distancia entre la rejilla fuente y de fase es 106 cm, mientras que la distancia entre la rejilla de fase y la analizadora es de 51 cm. Para más información ver [Scherer et al., 2014]. En la Figura 2.3 se puede ver la foto del montaje experimental utilizado.

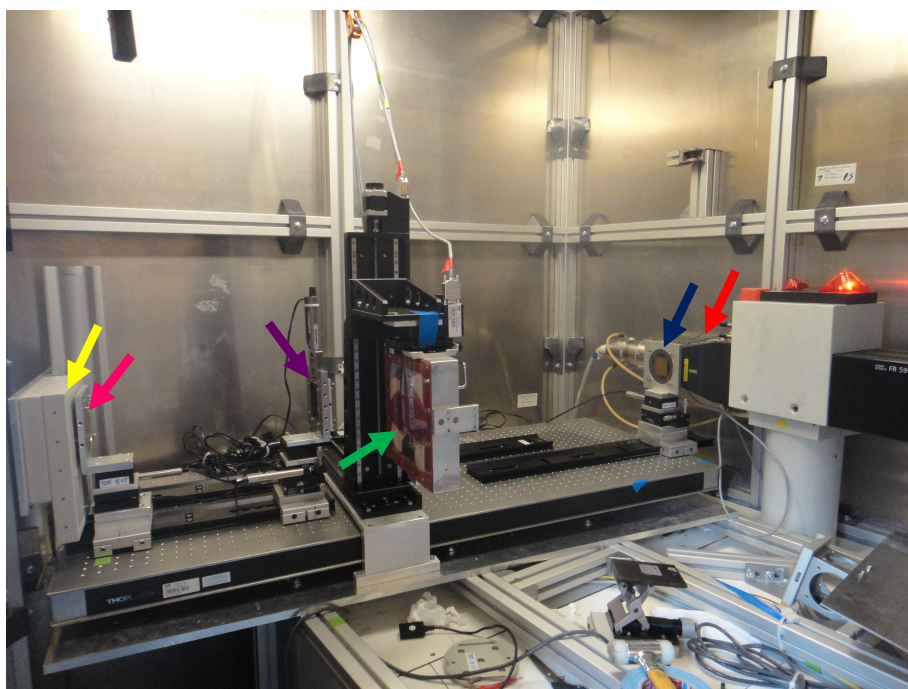


Figura 2.3: Foto del montaje experimental de *Technische Universität München* (TUM). Rojo: Fuente de rayos-X, Azul: Rejilla G0, Verde: Recipiente para la muestra, Morado: Rejilla G1, Rosa: Rejilla G2 y Amarillo: Sensor/detector.

El montaje había sido optimizado para tejidos mamarios, los cuales son similares en atenuación a los productos alimenticios, por su contenido en agua. La configuración del montaje no fue cambiada para evitar un extenso estudio de optimización. Idealmente, debería haber sido optimizado para cada uno de los

siete productos de los que se tomaron imágenes. La energía efectiva fue fijada a 25 keV y cada imagen fue tomada desde un sólo ángulo (una única proyección) y la rejilla movable fue desplazada 9 veces.

El tiempo de integración, t_{int} , fue el único parámetro que se cambió mientras se tomaban las imágenes. La mayor parte del tiempo estaba fijado a dos segundos, pero para dos de los productos (pavo y bistec con objetos de 4x4x4 mm) fue cambiado a un segundo. En la Figura 2.4 podemos observar que cuando el tiempo de exposición o integración es demasiado pequeño, 0.2 segundos, hay mucho ruido presente en la imagen. En contraposición, cuanto mayor es t_{int} , más artefactos aparecen en la imagen como, por ejemplo, fase envuelta, la cual se aprecia cuando t_{int} es dos segundos. Por tanto, es necesario encontrar un compromiso entre ruido y artefactos, por ello se elige el tiempo de exposición igual a un segundo para estos dos productos. No obstante, en general, dos segundos parecía ser un valor adecuado para el tiempo de exposición, es decir, un buen compromiso de cara a que las imágenes no incluyesen demasiado ruido o artefactos.

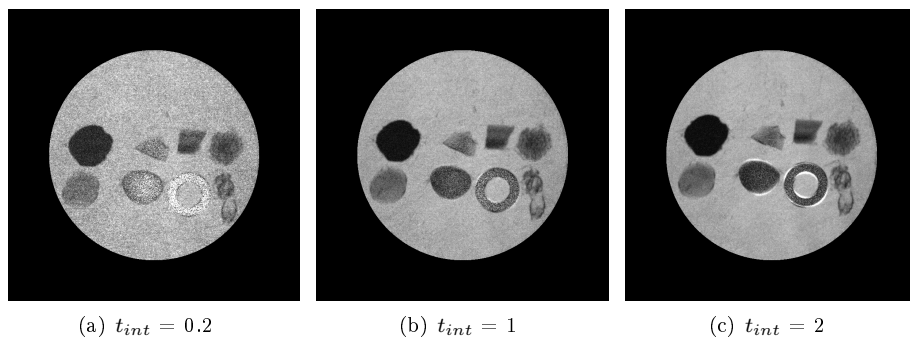


Figura 2.4: Imágenes de absorción de pavo para distintos tiempos de integración.

2.2 Materiales

Los datos constan de tres modalidades de imagen de comida con objetos extraños presentes en ella. Estos datos fueron adquiridos personalmente en *Technische Universität München*.

La adquisición de datos es un proceso lento, debido tanto al hecho de que las muestras tienen que ser preparadas manualmente en el recipiente como a que,

durante la toma de la imagen, el movimiento físico de la rejilla que analiza el patrón de interferencia es lento. Por esta razón, la opción de sintetizar datos fue considerada, pero desgraciadamente no es posible debido a que no se puede predecir la dispersión de los distintos materiales. Es decir, el tamaño de los objetos extraños en la modalidad de campo oscuro no puede ser pronosticado, ya que depende de cómo esté posicionado el objeto o incluso del tipo de madera, por ejemplo.

2.2.1 Comida

Para reproducir la variación de productos dentro de la industria alimenticia, ha sido elegida una gama de productos de homogeneidad variada para la toma de imágenes. El queso en lonchas representa un producto completamente homogéneo, mientras que el pan de centeno con semillas simboliza un producto no homogéneo. Se espera que los modelos para los productos homogéneos sean más sencillos.

Además, hay que elegir productos con distintas propiedades de dispersión, para cubrir diversas diferencias de contraste entre comida y objetos. El pan de trigo contiene harina, la cual es altamente dispersiva debido a su micro estructura refinada. Consecuentemente, se espera que los objetos extraños no resalten muy bien en la modalidad de campo oscuro para el pan de trigo, mientras que serán apreciados con mayor claridad en el pan de centeno, tal y como se puede ver en la Figura 2.5. Dicho de otra manera, cuanto más difieran las propiedades de dispersión de comida y objetos extraños, más fácil de identificar serán los objetos.

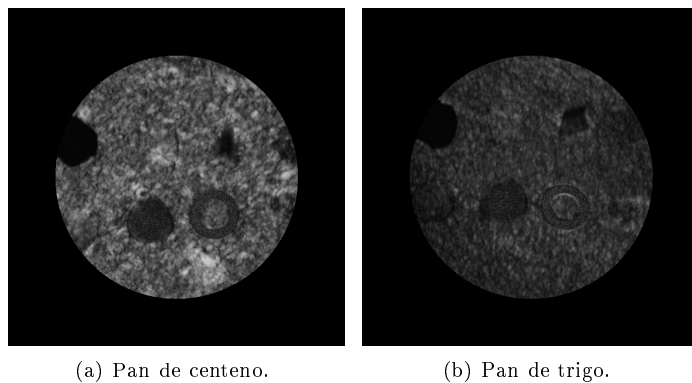


Figura 2.5: Diferentes contrastes de intensidad en campo oscuro.

Asimismo, como este proyecto forma parte de un proyecto más amplio, el proyecto NEXIM ¹, algunos de los productos alimenticios escogidos son aquellos producidos por las empresas colaboradoras en este estudio. La carne picada, el pan, el pollo y los productos lácteos son de interés para nuestros colaboradores industriales.

Finalmente, los productos seleccionados para escanear fueron carne picada, pan de trigo y de centeno, pavo (similar al pollo), queso, filete de ternera y salami. Todos los productos han sido modelados, pero sólo algunos van a ser tratados en esta memoria. El pavo y el filete son similares, por lo que solamente se tratará en profundidad el pavo. Salami y carne picada también son parecidos, ambos contienen tanto carne como grasa, entonces el producto que será tratado será el salami. El pan de centeno es elegido en lugar del pan de trigo, por ser mucho más importante para la industria danesa. De hecho, no merece la pena utilizar esta nueva técnica de escaneado para el pan de trigo, debido a la alta dispersión que éste presenta. También es más interesante el pan de centeno por ser el producto menos homogéneo en textura.

2.2.2 Objetos extraños

Los objetos ajenos son seleccionados para cubrir distintas propiedades de absorción, refracción y dispersión. Además, se tienen en cuenta tanto los objetos sugeridos por los colaboradores industriales de NEXIM (véase la sección 1.2) como los resultados de un estudio japonés [Takashi, 2009].

La Figura 2.6 muestra los resultados del estudio japonés, donde se presenta la frecuencia con la que los consumidores encontraban un contaminante concreto en comida contaminada, y la dificultad de detección de cada contaminante con rayos-X convencionales. Tal y como se ha dicho anteriormente, las técnicas de rayos-X convencionales miden las propiedades de absorción de los materiales, en otras palabras, los objetos serán más difícil de encontrar cuanto más parecida sea su densidad a la de la comida, la cual es cercana a la del agua ($1,000.00 \text{ kg/m}^3$). Por ello, el papel y otros materiales orgánicos son particularmente difíciles de encontrar.

¹<http://nexim.nbi.ku.dk/>

Contaminant	Percentage [%]	Difficulty detecting by X-ray	Contaminant	Percentage [%]	Difficulty detecting by X-ray
Insects	24.5	Difficult	Stone and sand	3.0	Easy
Bone (calcified only)	15.2	Medium	Paper, threads etc.	2.1	Difficult
Unclear	14.1	N/A	Vinyl	2.0	Difficult
Metal piece	7.3	Easy	Fly	1.8	Difficult
Hairs	6.6	Difficult	Wood chips	1.5	Difficult
Needles, wires etc.	6.5	Easy	Blade chips	1.2	Medium
Plastic and rubber	5.3	Medium	Staples	1.0	Easy
Glass fragments	3.9	Medium	Rat excrement	0.9	Difficult
Cockroach	3.1	Difficult			

Figura 2.6: Resultados de un estudio japonés sobre quejas de clientes relacionadas con contaminantes [Takashi, 2009].

Finalmente, se eligen dos tipos principales de objetos:

1. Aquellos que son fáciles de encontrar en la modalidad de absorción, como metal, piedra o cristal.
2. Aquellos que no son visibles en la modalida de absorción, como los que están compuestos por materia orgánica. En concreto, insectos, madera, goma y dos tipos de plástico (blando y duro).

En la Figura 2.7 se muestran los objetos extraños seleccionados. Los tamaños elegidos varían entre 2x2x2 y 4x4x4 mm aproximadamente, ya que se considera que objetos más pequeños no son peligrosos. En la Tabla 2.1 se muestra el espesor aproximado de los objetos, medido en la dirección en la cual son atravesados por los rayos-X.

Tipo	Espesor (mm)		
	2	3	4
1.Vidrio	2	3	5
2.Metal	0.5	1	2
3.Madera	2	4	6
4.Insectos	2	3	5
5.Plástico duro	2	3	6
6.Plástico blando	2	3	5
7.Goma	2	3	4
8.Piedras	3	4	6

Tabla 2.1: Espesor de los objetos en la dirección en que son atravesados por los rayos-X.



Figura 2.7: Objetos extraños seleccionados, con tamaños aproximados de 2x2x2, 3x3x3 y 4x4x4 mm, de izquierda a derecha.

Resumiendo, contamos con tres imágenes para cada uno de los siete productos alimenticios, cada imagen contiene ocho objetos extraños del mismo tamaño aproximado, ver Figura 2.8 como ejemplo.



Figura 2.8: Pavo con objetos extraños de distintos tamaños.

2.3 Métodos

En esta sección, se tratarán los métodos usados para el preprocesamiento y la adquisición de atributos a partir de las imágenes. Además, se hablará de las técnicas utilizadas para el entrenamiento de los clasificadores.

2.3.1 Preprocesado

En montajes reales en la industria, el fondo de la imagen debe ser eliminado previamente. Adicionalmente, las características del envase y el tamaño y forma de los productos cortados deberá ser modelado para cada industria y producto específico.

Este proyecto no se centra únicamente en un producto concreto, por lo que los productos seleccionados son escaneados sin incluir fondo y, en el caso del queso en lonchas, éstas son apiladas una encima de otra. Así, no es necesario este tipo de preprocesado.

De todas formas, sí que se necesita algo de preprocesado, pero el preprocesado realizado en las imágenes de este proyecto no sería necesario en una implementación en una cinta transportadora. Algunos píxeles necesitan ser eliminados, como aquellos pertenecientes a la cinta usada para fijar las muestras al recipiente y que no se escurrieran. Esto es resuelto con simple enmascaramiento.

Debido a las restricciones temporales de escaneo, no se pudieron obtener datos

de comida sin objetos extraños. Por tanto, es necesario etiquetar cada imagen en regiones pertenecientes a comida o a objetos extraños. Esto es usado tanto para el entrenamiento como para los resultados de detección para ambos conjuntos, validación y test.

2.3.2 Atributos

Para cada observación/píxel de la imagen, existen tres atributos espectrales directamente disponibles, correspondientes a las intensidades de cada píxel en cada una de las tres modalidades. No obstante, puede que sea más fácil discernir comida de objetos extraños si añadimos más atributos. Por ello, también se toma en consideración información espacial, en concreto, atributos de textura.

2.3.2.1 Atributos de textura

Existen diversas maneras de analizar la textura de una imagen. Algunas técnicas se centran en estadísticas de primer, segundo o incluso mayor grado [Cartensen, 1992], mientras que también se pueden utilizar otras técnicas más complejas, como los atributos wavelet [Arivazhagan and Ganesan, 2003], los textures [Leung and Malik, 1999] o los atributos básicos de imagen (BIF) [Crosier and Griffin, 2008].

Una de las razones que hace que los atributos básicos de imagen, *Basic Image Features*, sea un método popular, es el hecho de que permite la clasificación por texturas sin necesidad de optimizar parámetros para cada conjunto específico de datos [Crosier and Griffin, 2010]. Por esta razón, BIF es la técnica usada para realizar análisis de texturas durante este proyecto.

Basic Image Features (BIF) proporcionan un vector de respuesta a siete tipos de estructura local cualitativamente distintos. Esto se consigue aplicando a la imagen un conjunto de seis filtros derivativos Gaussianos, para un valor de escala de imagen concreto σ . Además, se tiene que fijar el parámetro de alisado ϵ , el cual determina el punto a partir del cual una superficie puede ser considerada uniforme (primer tipo de estructura local, a la izquierda de la Figura 2.9). En la Figura 2.9 se observan las distintas estructuras de imagen que son capturadas por esta técnica [Crosier and Griffin, 2010].



Figura 2.9: Estereotipos de área capturados por Basic Image Features (BIF) (Adaptada de [Crosier and Griffin, 2010]).

2.3.2.2 Regresión/Selección de atributos (FS)

En este caso de prueba concreto, los atributos básicos de imagen son calculados para tres escalas distintas ($\sigma = 1, 5, 10$), por lo que el número total de atributos disponibles será

1. Tres atributos de intensidad, correspondientes a cada una de las tres modalidades (absorción, campo oscuro y contraste de fase).
2. Siete atributos de textura por modalidad y escala, es decir, un total de 63 atributos de textura ($7 \text{ estructuras} \times 3 \text{ escalas} \times 3 \text{ modalidades}$).

Existen numerosos métodos para realizar regresión o selección de atributos

1. **Lasso** selecciona los atributos más relevantes imponiendo restricciones a la norma $|L|_1$ de los parámetros.
2. **Ridge o regularización de Tijonov** regulariza los atributos imponiendo restricciones a la norma $|L|_2$ de los parámetros.
3. **Elastic net** combina regularización con selección de atributos, permitiendo estimaciones robustas cuando el número de observaciones en el conjunto de datos es escaso.

El número de observaciones (píxeles) es muy alto comparado con el número de atributos, por lo que no se necesita realizar una selección de atributos. Para Lasso y Ridge sólo hay que optimizar un parámetro, mientras que para Elastic net hay que optimizar dos. Por simplicidad y adecuación al conjunto de datos, se elige la regresión Ridge.

Ridge minimiza la norma euclídea entre los datos y la línea de regresión ajustada a éstos, $|L|_2$; simultáneamente disminuye el impacto negativo de la colinealidad entre atributos. Por tanto, evita la cancelación entre atributos correlados que son importantes a la hora de predecir el valor de salida, mientras realiza la estimación de mínimos cuadrados en el contexto de regresión [McDonald, 2009].

2.3.3 Métodos de entrenamiento y clasificación

La tarea de clasificar cada píxel como comida u objeto requiere haber entrenado con anterioridad un modelo. Un gran número de métodos pueden ser utilizados para entrenamiento y clasificación, en esta sección se describen dos de los métodos más simples, pero más adecuados para este caso de prueba.

2.3.3.1 GMM

Un modelo de mezcla es un modelo probabilístico que representa la presencia de una o más poblaciones dentro de una población total. En particular, un modelo mezcla de Gaussianas (GMM) ajusta un número de Gaussianas preespecificado al conjunto de datos. A pesar de que el número de poblaciones sí que tiene que ser especificado, no hay necesidad de aportar más información sobre estas poblaciones, sólo las observaciones con sus atributos. En otras palabras, el modelo de aprendizaje es no supervisado. Además, puede haber cualquier número de atributos/dimensiones.

La distribución Gaussiana multidimensional ($k = n$ dimensiones/atributos) puede ser formulada de la siguiente manera

$$\phi(x|\mu, \Sigma) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp^{-1/2(x-\mu)^T \Sigma^{-1}(x-\mu)} \quad (2.1)$$

donde μ es el vector media y Σ es la matriz de covarianzas de la distribución Gaussiana. Con una probabilidad ϕ , la distribución tomara como valor el vector n -dimensional x .

La distribución mezcla de L Gaussianas es

$$p(x) = \sum_{l=1}^L p_l \phi(x|\mu_l, \Sigma_l) \quad (2.2)$$

donde p_l es la probabilidad a priori de cada Gaussiana perteneciente a la mezcla.

Para ajustar estas Gaussianas a los datos multidimensionales, se utiliza el algoritmo esperanza-maximización (EM) [Dempster et al., 1977]. Las Gaussianas son inicializadas con media aleatoria y con la matriz identidad como matriz

de covarianzas. Los pasos esperanza (E), que asigna un peso a cada grupo, y maximización (M), que calcula la nueva media y matriz de covarianzas de cada grupo, son alternados hasta que se encuentran los parámetros de máxima verosimilitud, el mejor ajuste posible.

Ventajas de GMM

1. El ruido presente en imágenes es típicamente Gaussiano en intensidad, es decir, es ruido blanco.
2. La intensidad de la comida puede ser considerada Gaussiana con la media fijada por el agua, el componente principal de la comida.
3. Sólo la comida tiene que ser representada en el conjunto de entrenamiento, por lo que cualquier tipo de objeto extraño podría ser encontrado, incluso objetos desconocidos.

Desventajas de GMM

1. El número de grupos o subpoblaciones tiene que ser supuesto, preespecificado. Esta tarea no es nada fácil, ya que los datos no pueden ser visualizados cuando hay más de tres atributos. Además, si este número no es correcto, el modelo será altamente inadecuado.
2. Es asumido que la distribución de probabilidad es normal. Probablemente la distribución de la comida sea parecida a, pero no exactamente, una Gaussiana, por lo que los resultados no serán perfectos.

La distancia de Mahalanobis, $D_M(x)$ es la medida que ha sido seleccionada en este proyecto para calcular el parecido entre los píxeles a clasificar y el modelo Gaussiano ajustado a los píxeles de comida.

$$D_M(x) = \sqrt{(x - \mu)^T \Sigma^{-1} (x - \mu)} \quad (2.3)$$

donde $x = (x_1, x_2, x_3, \dots, x_n)^T$ es el vector de atributos para cada una de las N observaciones/píxeles y μ y Σ son, respectivamente, el vector media y la matriz de covarianzas del modelo de comida.

Las ventajas de la distancia de Mahalanobis sobre la distancia Euclídea son la invarianza de escala y el hecho de que la distancia de Mahalanobis considera las correlaciones entre el conjunto de datos [Bose, 1993].

Tras calcular la distancia de Mahalanobis, se aplicará un valor umbral fijo (previamente optimizado) para segmentar la imagen, es decir, clasificar cada píxel como perteneciente a una categoría, comida u objeto extraño.

2.3.3.2 SVM

La máquina de vectores de soporte (SVM) es un modelo de aprendizaje supervisado. Un conjunto de entrenamiento, previamente etiquetado, es introducido al algoritmo de entrenamiento de SVM, el cual encuentra una separación entre las clases etiquetadas. En algunos casos, no es posible encontrar una solución perfecta o, si ésta existe, puede que no interese porque el modelo se haya sobreadjustado a los datos de entrenamiento y, por tanto, no sea capaz de generalizar. Por consiguiente, SVM permite que algunas observaciones sean clasificadas erróneamente, a cambio de obtener un modelo más simple. Esto es ilustrado en la Figura 2.10, donde se encuentra una separación entre clases, aunque algunas observaciones no han sido clasificadas correctamente en la Figura 2.10(b). La frontera que divide las clases (comida y objetos extraños) puede ser elegida para que sea tan simple como una función lineal o tan compleja como decida el usuario, ya que existe la posibilidad de introducir manualmente una función no preespecificada [Hastie et al., 2008].

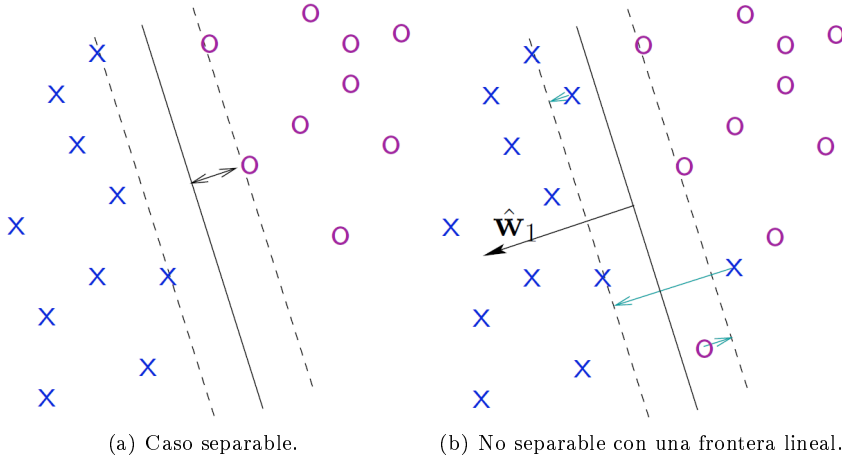


Figura 2.10: Clasificadores de vectores de soporte.

El problema es cuadrático con restricciones de desigualdad lineales. En otras palabras, es un problema de optimización convexo, esto es, de minimización.

$$\begin{aligned}
 & \underset{\beta, \beta_0}{\operatorname{argmin}} \frac{1}{2} \|\beta\|^2 + C \sum_{i=1}^n \xi_i \\
 & \text{subject to } \xi_i \geq 0, y_i(x_i^T \beta + \beta_0) \geq 1 - \xi_i \forall i
 \end{aligned} \tag{2.4}$$

donde el parámetro de penalización C es añadido para el caso no separable; el caso separable se corresponde $C = \infty$ [Hastie et al., 2008]. Cuanto mayor es C , mayor es la penalización para los errores y puede que exista sobreajuste hacia los datos de entrenamiento. Mientras que si es demasiado pequeño, tendremos subajuste [Khan, 2008].

Para este caso de prueba concreto, se utilizará una frontera lineal por simplicidad, y se permitirá el solapamiento entre clases hasta cierto punto.

Al igual que en GMM, SVM también puede manejar cualquier número de atributos/dimensiones. De hecho, conforme crece el número de atributos, aumenta la probabilidad de una correcta separación entre clases.

Ventajas de SVM

1. No se hace ninguna suposición sobre el tipo de distribución de probabilidad. Aunque la comida puede ser aproximada por una mezcla de Gaussianas, la clasificación podría ser más precisa sin una suposición previa sobre el tipo exacto de distribución.

Desventajas de SVM

1. Para este caso concreto, el tiempo de cálculo es extremadamente alto, debido a que el número de observaciones es elevado. En otros casos, SVM presenta la ventaja, frente a otros métodos, de poder manejar muchos atributos, aún cuando el número de observaciones disponibles no es demasiado grande.
2. Es necesario que el conjunto de datos de entrenamiento contenga píxeles de todas las clases posibles. Como consecuencia, los objetos extraños que estén compuestos de materiales con propiedades diferentes a aquellos representados en el conjunto de entrenamiento, probablemente no sean encontrados. En contraposición, GMM podía encontrar cualquier tipo de objeto extraño.

Resultados

Esta sección se centra en crear y comparar distintos modelos. Para ello, se utilizará un solo producto, pavo. Varios modelos Gaussianos, con diferente número de atributos, serán ajustados a los píxeles de comida. Luego, estos modelos serán optimizados, y se compararán los resultados de la clasificación basada en píxeles para determinar el conjunto óptimo de modalidades de imagen para este caso de prueba, además se examinará si merece la pena utilizar atributos de textura. Más tarde, el modelo que mejores resultados proporcione se comparará con el mejor modelo de los dos que han sido creados únicamente con la modalidad de absorción, uno de estos dos modelos contendrá análisis de textura y el otro no. Por último, se comparará el funcionamiento de SVM y GMM para un modelo concreto.

Para determinar la robustez del modelo óptimo, se analizarán tres productos más. ¿Variará el conjunto óptimo de atributos de un producto a otro, dependiendo de las propiedades de absorción, refracción y dispersión del producto alimenticio en cuestión? ¿Cuánto mejor es el modelo óptimo, creado con estas nuevas modalidades de imagen, en comparación con el modelo óptimo que obtendríamos utilizando sólo la absorción?

3.1 Conjunto de datos principal: Pavo

Para detectar objetos extraños, es necesario entrenar un modelo. A continuación, los parámetros del modelo necesitan ser optimizados y, finalmente, se necesita evaluar la efectividad del modelo. Por tanto, el número total de observaciones/píxeles contenido en las tres imágenes, correspondientes a los tres tamaños de objetos, es dividido en tres conjuntos: entrenamiento, validación y test. Los píxeles de comida pertenecientes al conjunto de entrenamiento son utilizados para el ajuste de los modelos de comida, mientras que los píxeles que forman parte del conjunto de validación son utilizados para optimizar los parámetros de los modelos. Para comparar el rendimiento de los distintos modelos se utiliza el conjunto de test.

3.1.1 Modelos Gaussianos

Los siguientes modelos Gaussianos, de distintas dimensiones, son ajustados al producto alimenticio.

Modelo	Nº Atrib.	Absorción	Cont. de fase	Campo oscuro	BIF
1	1	Sí	No	No	No
2	1	No	Sí	No	No
3	1	No	No	Sí	No
4	2	Sí	Sí	No	No
5	2	Sí	No	Sí	No
6	2	No	Sí	Sí	No
7	3	Sí	Sí	Sí	No
8	22	Sí	No	No	Sí
9	22	No	Sí	No	Sí
10	22	No	No	Sí	Sí
11	44	Sí	Sí	No	Sí
12	44	Sí	No	Sí	Sí
13	44	No	Sí	Sí	Sí
14	66	Sí	Sí	Sí	Sí

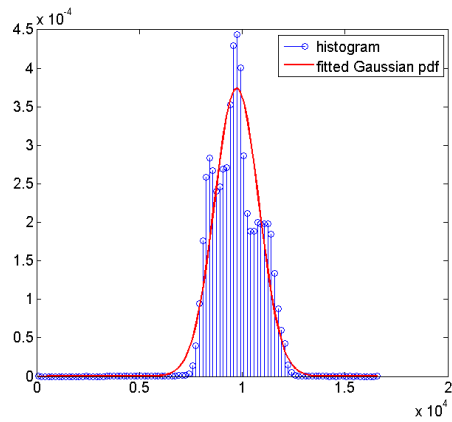
Tabla 3.1: Diferentes modelos Gaussianos, incluyendo un número variable de atributos y de modalidades.

Los parámetros que tienen que ser optimizados cuando se emplean modelos Gaussianos son los siguientes:

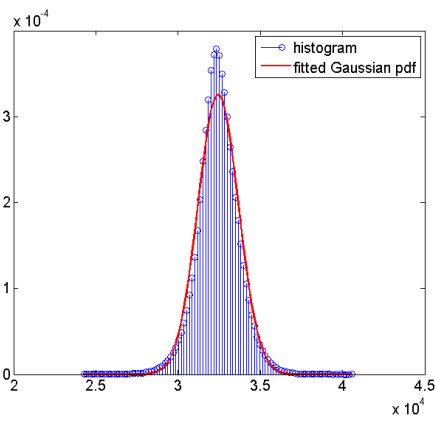
1. D , el número de Gaussianas a ajustar a los datos de comida de entrenamiento.
2. Th , el umbral que va a ser aplicado a la distancia de Mahalanobis para decidir a qué categoría pertenece cada píxel.

Para los siete modelos que incorporan los atributos de textura, también tendrá que ser ajustado el parámetro de la regresión Ridge, k_{ridge} .

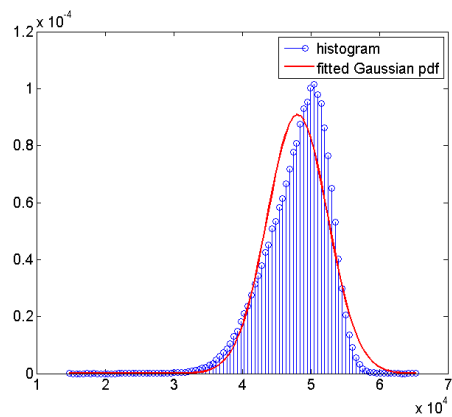
Las estadísticas de primer grado, como los histogramas de la Figura 3.1, son útiles para determinar cuántas Gaussianas han de ser ajustadas a cada tipo de producto alimenticio. Simples diagramas de dispersión, como los representados en la Figura 3.2, también proporcionan información sobre el número de Gaussianas que debería ser ajustado a los datos de comida.



(a) Absorción

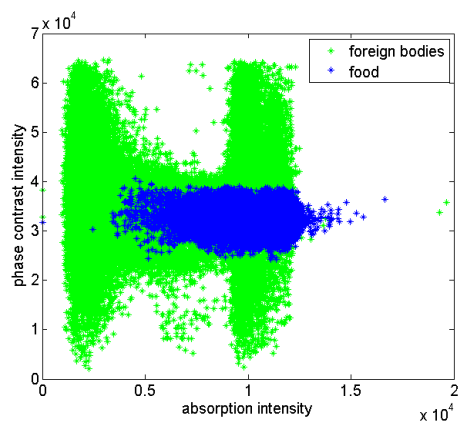


(b) Contraste de fase

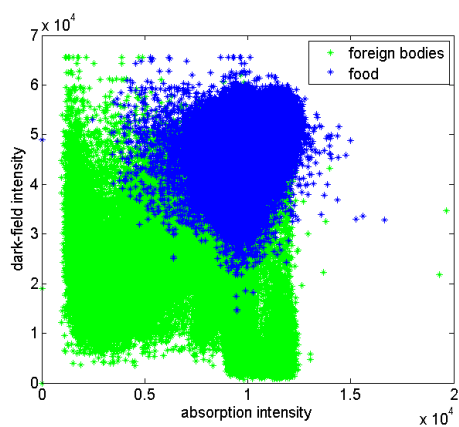


(c) Campo oscuro

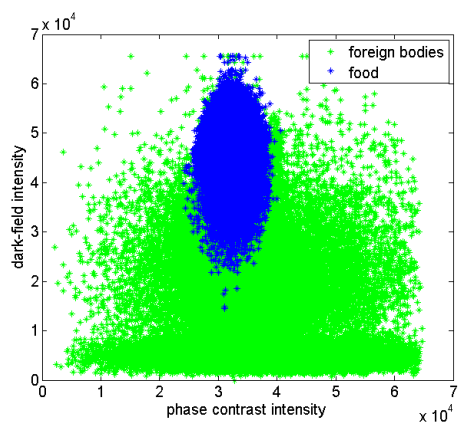
Figura 3.1: Histogramas y Gaussianas ajustadas a las intensidades del pavo en las distintas modalidades.



(a) Absorción y contraste de fase.



(b) Absorción y campo oscuro.



(c) Contraste de fase y campo oscuro.

Figura 3.2: Diagramas de dispersión 2D para las intensidades del pavo.

Observando las Figuras 3.1 y 3.2, se decide ajustar una única Gaussiana multidimensional a los datos de entrenamiento de comida, es decir, el parámetro D será uno. Por simplicidad, se utilizará $D = 1$ para los catorce modelos, a pesar de que algunos modelos incluyen más atributos, no sólo los espectrales. Adicionalmente, en las Figuras 3.1 y 3.2 se puede ver que los datos de comida se aproximan a una distribución normal, ya que existe una mayor densidad de observaciones en el centro de la distribución y una menor densidad conforme nos alejamos del interior de la distribución.

El parámetro k_{ridge} es optimizado utilizando el conjunto de validación, se elige como k_{ridge} óptimo aquel k_{ridge} que minimiza el error de validación, definido como

$$error = \sqrt{\left(\frac{FN}{N}\right)^2 + \left(\frac{FP}{P}\right)^2} \quad (3.1)$$

donde

1. FN es el número de falsos negativos en el conjunto de entrenamiento, esto es el número de píxeles clasificados como comida, pero que realmente pertenecen a la clase positiva (objetos extraños).
2. FP es el número de falsos positivos, es decir, el número de píxeles clasificados como positivos, pero que en realidad forman parte de la clase negativa (comida).
3. N es el número de píxeles que pertenecen a la clase negativa, es decir, el número de píxeles de comida que hay en el conjunto de validación.
4. P es el número de píxeles que forman parte de la clase positiva, es decir, el número de píxeles en los datos de validación que se corresponden con objetos extraños.

En la Figura 3.3, se muestran los errores de validación y entrenamiento para uno de los modelos, aquel que incluye todos los atributos posibles.

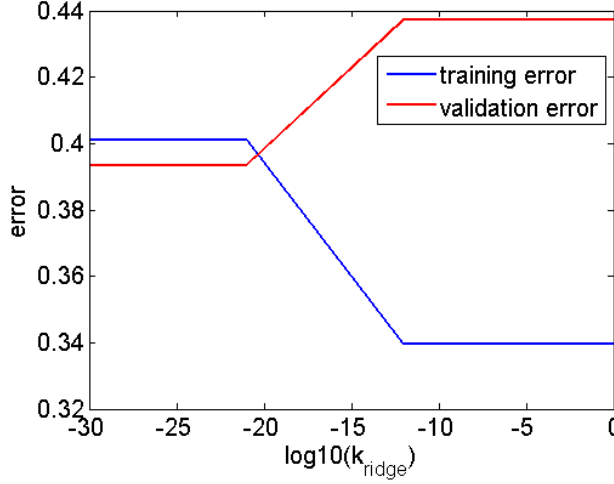
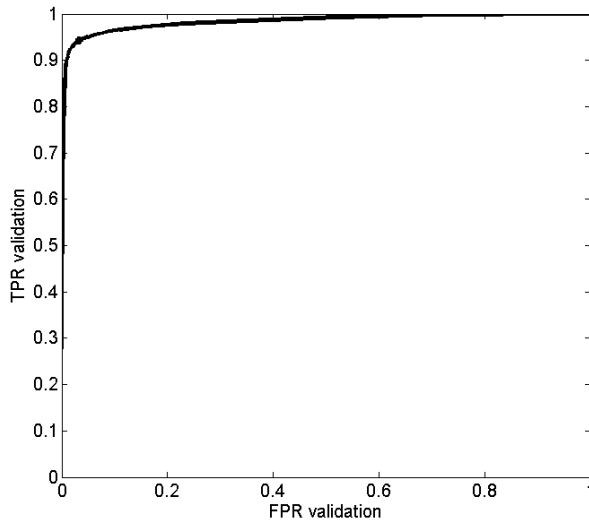


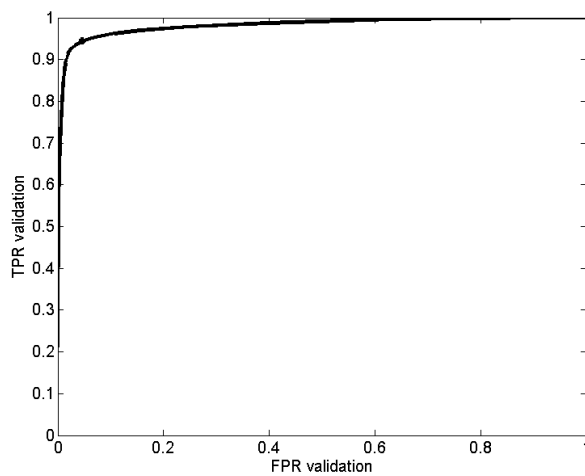
Figura 3.3: Errores de entrenamiento y validación para el modelo de pavo número 14.

Tal y como se puede ver en la Figura 3.3, el error de validación no es muy sensible a los cambios del parámetro k_{ridge} . Esto se puede evaluar mejor mediante la observación de las curvas ROC. La característica operativa de un receptor (curva ROC) ilustra el rendimiento de un clasificador binario y se crea dibujando la sensibilidad o razón de verdaderos positivos (TPR) frente a la razón de falsos positivos (FPR), equivalente a 1-especificidad, para varios umbrales o puntos de corte (valor desde el cual se decide que un píxel concreto pertenece a la clase positiva, es decir, es un objeto extraño).

En la Figura 3.4, se ilustran las curvas ROC para el mejor y el peor caso de k_{ridge} , aquellos valores que proporcionan el menor y el mayor error de validación, respectivamente. El punto óptimo de la curva ROC, correspondiente con el umbral óptimo que debe ser aplicado a la distancia de Mahalanobis, es aquel que se encuentra más cerca del clasificador perfecto (TPR=1, FPR=0) en distancia euclídea d . Además, la distancia euclídea se puede utilizar para evaluar y comparar la efectividad o el rendimiento de los modelos.



(a) Mejor caso, optimizado ($k_{ridge} = 10^{-21}$), $TPR = 0.9437$, $FPR = 0.0337$ y $d = 0.0656$.



(b) Peor caso ($k_{ridge} = 10^{-12}$), $TPR = 0.9447$, $FPR = 0.0468$ y $d = 0.0725$.

Figura 3.4: Curva ROC para el modelo 14 de pavo con regresión Ridge.

La diferencia entre el modelo con k_{ridge} optimizado y el modelo para el peor caso de k_{ridge} es tan pequeña que no merece la pena optimizar este parámetro porque es costoso en lo que se refiere al tiempo. Puede que no merezca la pena utilizar la regresión Ridge en absoluto. Para evaluar esto, se muestra en la Figura 3.5 la curva ROC para el modelo 14 de pavo, pero sin incluir la regresión Ridge.

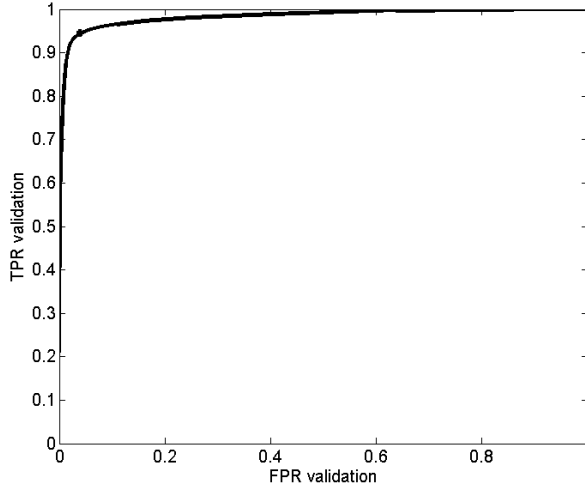
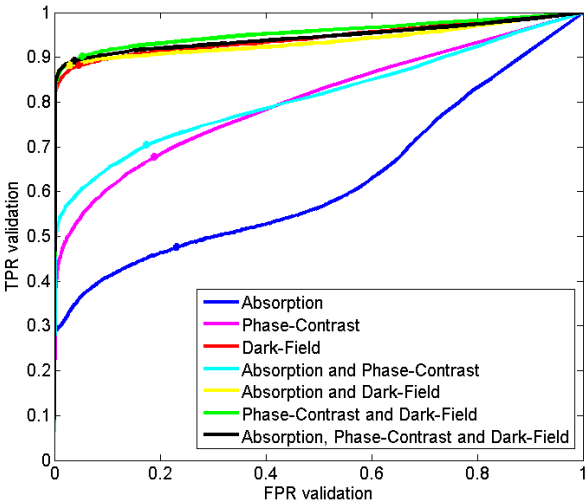


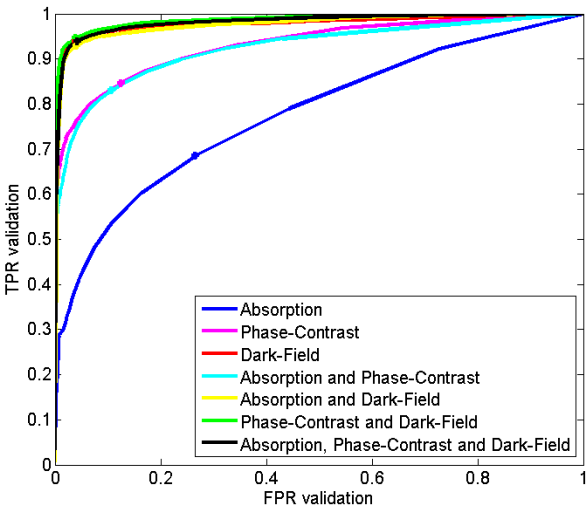
Figura 3.5: Curva ROC para el modelo 14 de pavo sin regresión Ridge.
 $TPR = 0.9448$, $FPR = 0.0393$ y $d = 0.0678$.

Al observar la Figura 3.5 se deduce que la regresión Ridge no produce una mejora significativa, ya que la distancia desde el punto óptimo al clasificador ideal es 0.0656 cuando se utiliza Ridge, comparando con 0.0678 cuando no se utiliza la regresión Ridge. Estas distancias son muy parecidas, por lo que la regresión Ridge no será usada de aquí en adelante para evitar un costoso estudio de optimización que no aporta beneficios significativos.

En la Figura 3.6, se han optimizado los umbrales para cada uno de los 14 modelos de comida. Los puntos óptimos son los puntos más cercanos de cada curva ROC al clasificador ideal ($TPR=1$, $FPR=0$), y están marcados en las gráficas.



(a) Modelos de pavo con atributos BIF.



(b) Modelos de pavo sin atributos BIF.

Figura 3.6: Curvas ROC para los modelos del pavo.

Mirando las curvas ROC, se puede comparar cualitativamente la efectividad de cada uno de los modelos. No obstante, la efectividad también se puede evaluar cuantitativamente, observando las distancias euclídeas entre cada uno de los 14 modelos óptimos y el clasificador ideal. Estas distancias se muestran en la Tabla 3.2, empezando por el modelo que proporciona el mejor rendimiento para una clasificación basada en píxeles.

Distance d	Model Nr.
0.0664	13
0.0720	10
0.0739	14
0.0811	12
0.1113	6
0.1135	7
0.1205	5
0.1260	3
0.1976	9
0.2003	11
0.3437	4
0.3741	2
0.4116	8
0.5723	1

Tabla 3.2: Distancias desde los modelos optimizados al clasificador ideal.

Tal y como se puede ver en la Tabla 3.2 y en la Figura 3.6, el rendimiento de los clasificadores mejora en gran medida tras añadir un análisis de textura, por lo que sí es beneficioso añadir estos atributos espaciales adicionales. Asimismo, el modelo óptimo para el pavo es el que incluye las modalidades de contraste de fase y campo oscuro y el peor modelo es aquel que incluye solamente la intensidad de la modalidad de absorción.

Lo más correcto es comparar el funcionamiento de los modelos sobre el conjunto de observaciones/píxeles de test. En la Tabla 3.3 se muestra la matriz de confusión (CM) para el modelo número 13 y para el mejor modelo de absorción de los dos disponibles (1 y 8, sin y con análisis de textura respectivamente). Además, se muestra la exactitud y la precisión en recuperación de información. Estas medidas están definidas a continuación.

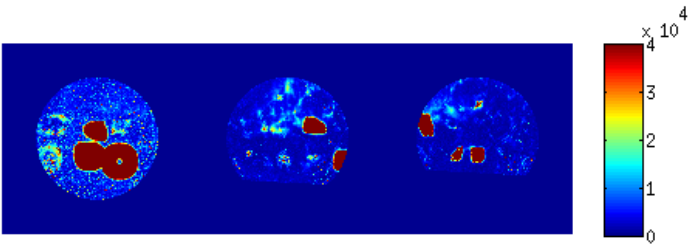
$$exactitud = \frac{TP + TN}{P + N} \quad (3.2)$$

$$precision = TP/P \tag{3.3}$$

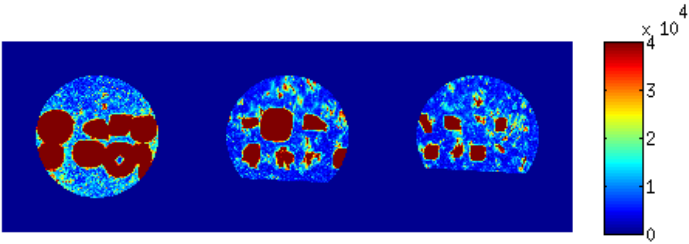
Modelo	N° Atrib.	TP	FP	TN	FN	Exac.	Prec.
13	44	3569	1211	38806	225	0.9407	0.9672
8	22	2624	10183	29834	1170	0.6916	0.7409

Tabla 3.3: Rendimiento de los modelos de pavo sobre el conjunto de test.

En la Figura 3.7, se observa la matriz de distancias para los dos modelos de la Tabla 3.3. El primer modelo es el modelo de pavo que clasifica mejor de entre los dos modelos que incluyen única y exclusivamente la modalidad de absorción, es decir, el mejor modelo de entre los modelos 1 y 8. El segundo modelo es el modelo de pavo que proporciona la mejor clasificación, de entre los 14 modelos disponibles. Además, en la Figura 3.8, se muestra la asignación final de clases (píxel a píxel) correspondiente a estos modelos.

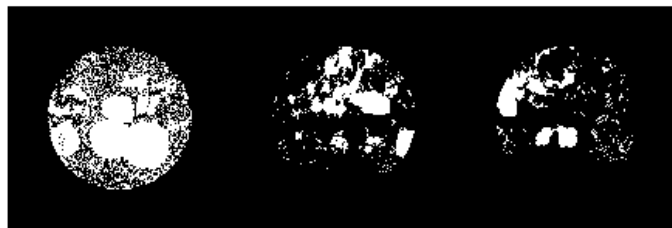


(a) Mejor modelo de absorción para el pavo, modelo número 8.



(b) Mejor modelo para el pavo, modelo número 13.

Figura 3.7: Mapas de distancias.



(a) Mejor modelo de absorción para el pavo, modelo número 8.



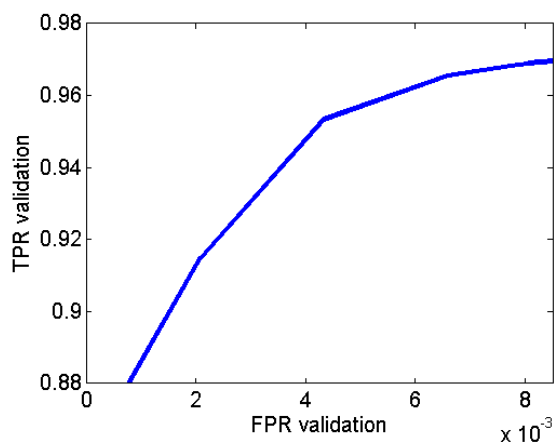
(b) Mejor modelo para el pavo, modelo número 13.

Figura 3.8: Clasificación resultante tras aplicar el umbral óptimo a la distancia de Mahalanobis.

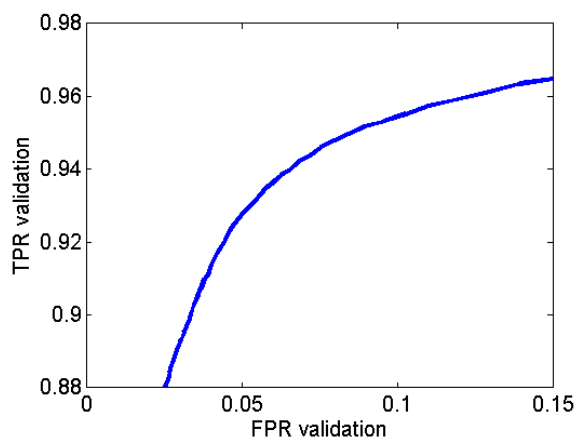
3.1.2 Modelo Gaussiano VS Máquina de Vectores de Soporte

Al igual que para GMM, el conjunto de datos es dividido en tres conjuntos: entrenamiento, validación y test. Pero a diferencia de GMM, el conjunto de entrenamiento no sólo contiene píxeles correspondientes a comida, sino que también cuenta con observaciones de los objetos extraños.

Se va a comparar el funcionamiento del modelo Gaussiano frente a SVM para el modelo número 10, el cual incluye la modalidad de campo oscuro, con sus atributos tanto espectrales como espaciales.



(a) Modelo SVM de pavo, $d = 0.0468$.



(b) Modelo GMM de pavo, $d = 0.0712$.

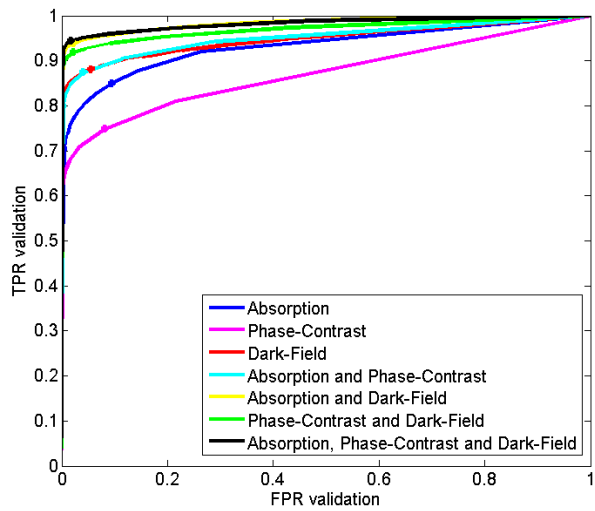
Figura 3.9: Curvas ROC para los modelos de campo oscuro de pavo, incluyendo atributos de textura.

Mirando la Figura SVM 3.9, se puede ver de forma cualitativa que el modelo SVM es un mejor clasificador. Además, también se puede ver la eficacia de SVM cuantitativamente a través de los valores de distancia euclídea al clasificador ideal d ($d = 0.0468$ para SVM y $d = 0.0712$ para GMM). Sin embargo, se tiene que tener en cuenta el hecho de que el rendimiento del modelo SVM disminuirá rápidamente si se introducen objetos extraños en el conjunto de test, diferentes de aquellos contenidos en el conjunto de entrenamiento. Además de que GMM puede detectar objetos sin que tengan que formar parte del conjunto de entrenamiento, también es un modelo mucho más rápido de entrenar.

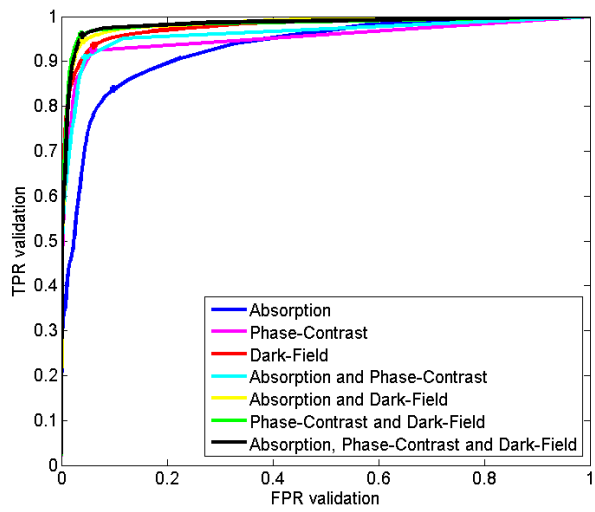
3.2 Conjunto de datos ampliado: Queso, Salami y Pan de centeno.

Los tres productos serán analizados a continuación utilizando modelos Gaussianos sin regresión Ridge, por las razones mencionadas en el apartado anterior.

Se obtienen los siguientes resultados para queso (Figura 3.10), salami (Figura 3.11) y pan de centeno (Figura 3.12). Gracias a las curvas ROC, se pueden comparar los distintos modelos de forma visual.

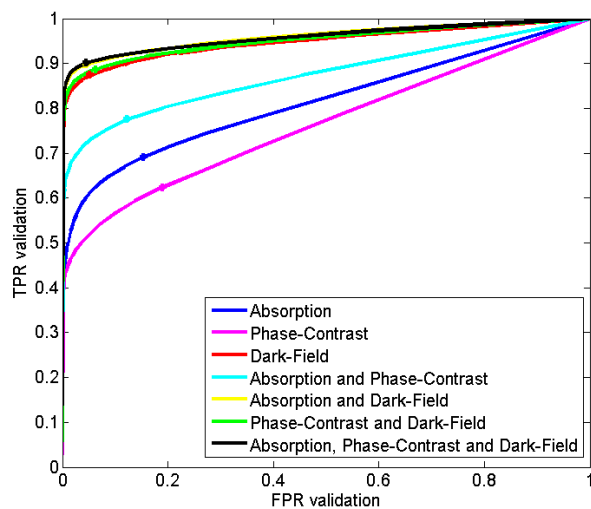


(a) Modelos de queso sin atributos BIF.

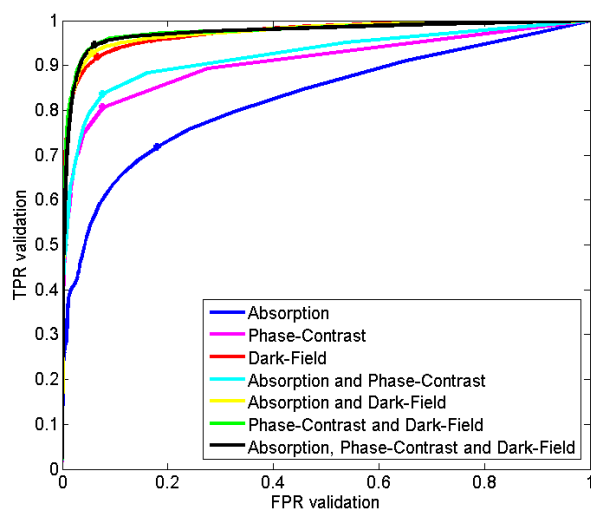


(b) Modelos de queso con atributos BIF.

Figura 3.10: Curvas ROC para los modelos de queso.

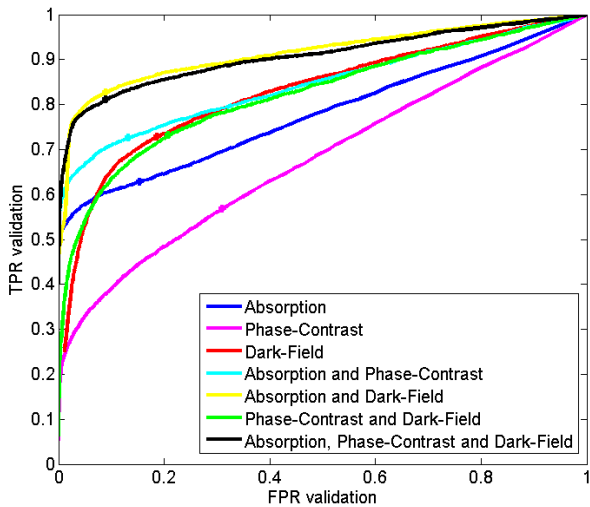


(a) Modelos de salami sin atributos BIF.

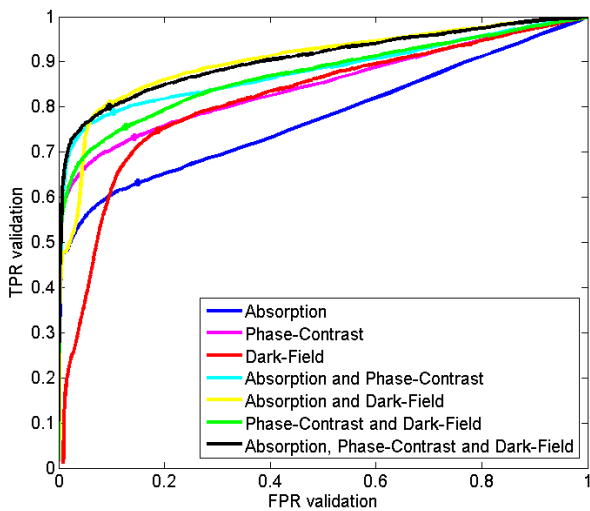


(b) Modelos de salami con atributos BIF.

Figura 3.11: Curvas ROC para los modelos de salami.



(a) Modelos de pan de centeno sin atributos BIF.



(b) Modelos de pan de centeno con atributos BIF.

Figura 3.12: Curvas ROC para los modelos de pan de centeno.

El funcionamiento de los distintos modelos puede ser evaluado cuantitativamente comparando la distancia euclídea d desde el punto óptimo al clasificador ideal. Estas distancias se muestran en la Tabla 3.4 para queso, salami y pan de centeno.

Queso		Salami		Pan de centeno	
Nº Modelo	d	Nº Modelo	d	Nº Modelo	d
13	0.0563	13	0.0814	5	0.1936
14	0.0576	14	0.0820	7	0.2081
7	0.0632	12	0.0915	12	0.2146
12	0.0668	10	0.1054	14	0.2215
5	0.0678	7	0.1083	11	0.2353
6	0.0854	5	0.1119	13	0.2747
10	0.0892	6	0.1286	9	0.3026
9	0.0978	3	0.1362	4	0.3037
11	0.1021	11	0.1801	10	0.3143
3	0.1271	9	0.2078	3	0.3282
4	0.1381	4	0.2551	6	0.3388
1	0.1794	8	0.3335	8	0.3969
8	0.1882	1	0.3445	1	0.4011
2	0.2694	2	0.4209	2	0.5304

Tabla 3.4: Distancias desde los modelos optimizados al clasificador ideal.

Tal y como se puede ver en la Tabla anterior, distintos modelos funcionan de manera distinta dependiendo del producto en cuestión. El modelo 13 es el mejor para queso y salami, pero no para el pan de centeno, siendo mejor el modelo 5 para este producto.

Igualmente, se observa una mejoría de detección cuando se añaden los atributos BIF a los modelos de queso y salami, pero parece que los atributos BIF no mejoran los modelos de pan de centeno.

De hecho, la razón de detección es bastante pobre para pan de centeno, esto podría ser debido al hecho de que las semillas son detectadas como objetos extraños. Para captar correctamente la textura del pan de centeno con sus semillas, se debería realizar un estudio de optimización extenso sobre las escalas y los parámetros de alisado. Puede que se obtengan mejores resultados una vez se hayan optimizado estos parámetros. Aún así, el hecho de incluir las modalidades de contraste de fase y campo oscuro proporciona una mejor detección de objetos extraños en el pan de centeno, que usando únicamente la modalidad de absorción.

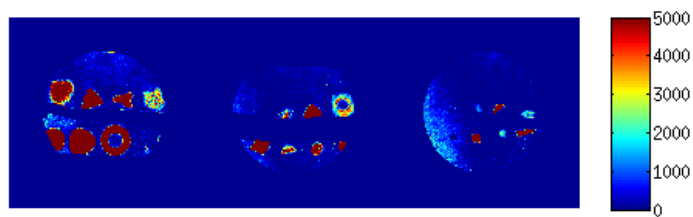
En la Tabla 3.5 se pueden comparar el rendimiento de los modelos sobre el conjunto de test.

Producto	Modelo	N° At.	TP	FP	TN	FN	Exact.	Prec.
Queso	13	44	7984	1268	32604	352	0.9578	0.9616
	1	1	7129	3494	30379	1189	0.8571	0.8890
Salami	13	44	6834	2125	37484	449	0.9383	0.9451
	8	22	5491	8682	30917	1803	0.7528	0.7764
Pan de centeno	12	44	2319	4582	43245	541	0.8108	0.8989
	8	22	1787	6934	40910	1056	0.6286	0.8424

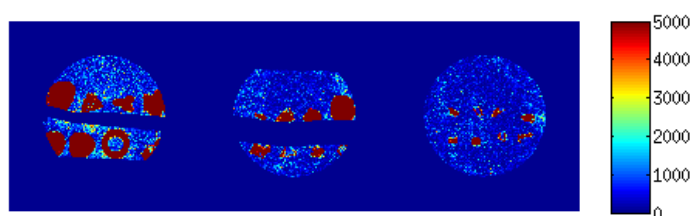
Tabla 3.5: Efectividad de los modelos sobre el conjunto de test.

Finalmente, se compara la clasificación píxel a píxel para el mejor de los dos modelos que contienen sólo la modalidad de absorción frente al modelo que clasifica mejor de los 14. Esto proporciona una representación visual de la mejora obtenida tras añadir estas nuevas modalidades de imagen: contraste de fase y campo oscuro.

En las Figuras 3.13, 3.15 y 3.17 se pueden observar los mapas de distancias para el modelo que proporciona la mejor clasificación frente al mejor modelo de los dos modelos que incluyen la modalidad de absorción solamente (modelos 1 y 8). En la Figura 3.14, 3.16 y 3.18 se muestran las decisiones finales de clasificación para cada píxel.



(a) Mejor modelo de absorción para el queso, modelo número 1.



(b) Mejor modelo para el queso, modelo número 13.

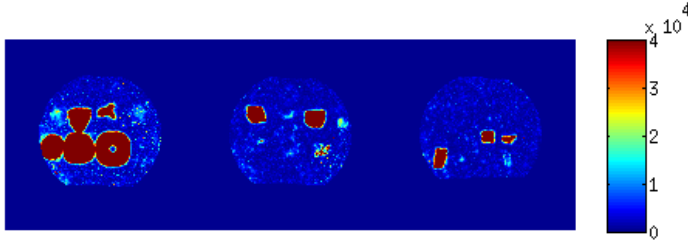
Figura 3.13: Mapa de distancias.

(a) Mejor modelo de absorción para el queso, modelo número 1

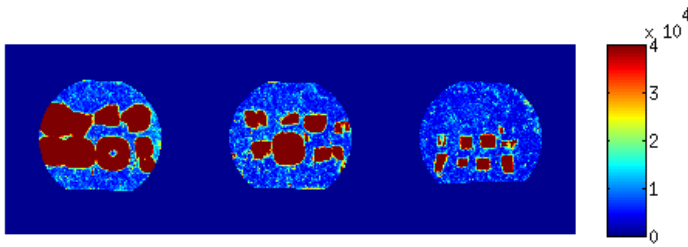


(b) Mejor modelo para el queso, modelo número 13.

Figura 3.14: Clasificación resultante tras aplicar el umbral óptimo a la distancia de Mahalanobis.

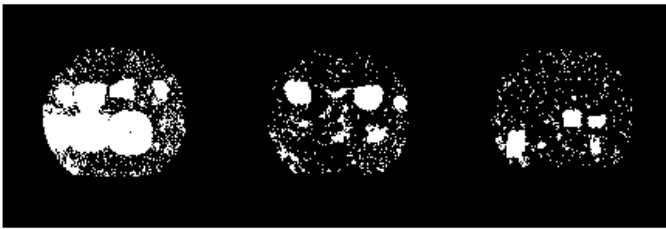


(a) Mejor modelo de absorción para el salami, modelo número 8.



(b) Mejor modelo para el salami, modelo número 13.

Figura 3.15: Mapa de distancias.

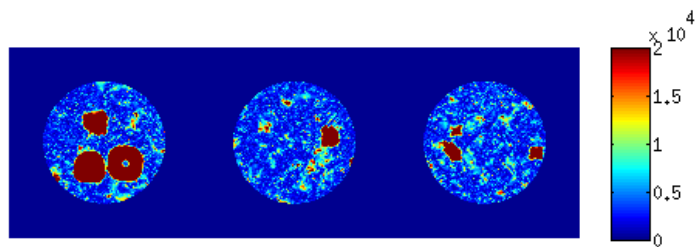


(a) Mejor modelo de absorción para el salami, modelo número 8.

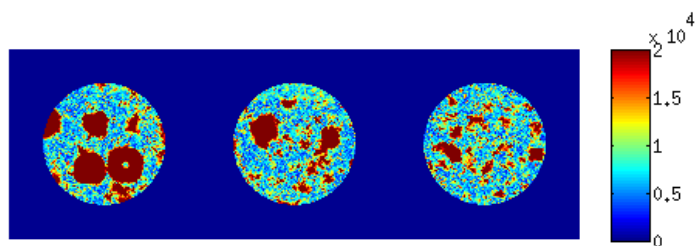


(b) Mejor modelo para el salami, modelo número 13.

Figura 3.16: Clasificación resultante tras aplicar el umbral óptimo a la distancia de Mahalanobis.



(a) Mejor modelo de absorción para pan de centeno, modelo número 8.



(b) Mejor modelo para pan de centeno, modelo número 12.

Figura 3.17: Mapa de distancias.

(a) Mejor modelo de absorción para pan de centeno, modelo número 8.



(b) Mejor modelo para pan de centeno, modelo número 12.

Figura 3.18: Clasificación resultante tras aplicar el umbral óptimo a la distancia de Mahalanobis.

CAPÍTULO 4

Conclusiones y trabajo futuro

El resultado más importante de este proyecto es la evidencia sobre una detección automática de objetos más eficiente tras añadir, a la modalidad de contraste de absorción, dos modalidades de imagen adicionales (contraste de fase y campo oscuro). Gracias a estas dos nuevas modalidades, podemos encontrar objetos orgánicos, lo cual era difícil utilizando rayos-X convencionales.

Los objetos orgánicos destacan mejor en la modalidad de campo oscuro, además en esta modalidad también se pueden ver los objetos no orgánicos. De hecho, la modalidad de absorción no añade mucha más información que la que ya contiene la modalidad de campo oscuro. No obstante, las modalidades de contraste de fase y de absorción pueden añadir algo de información extra a la modalidad de campo oscuro. Para el queso, la modalidad de absorción añade más información que la de fase, mientras que para los otros tres productos añade más información el contraste de fase. Si no han sido añadidos atributos de textura, los modelos que sólo incluyen la modalidad de absorción, suelen funcionar mejor que los que sólo incluyen contraste de fase. En lo que se refiere al análisis de textura, el cálculo de los atributos BIF merece la pena porque se obtiene una mejor detección tras añadir estos atributos espaciales. No obstante, si la textura del producto es compleja, como la del pan de centeno, se debe realizar previamente un extenso estudio para seleccionar los parámetros de escala y alisado que aportan mayor cantidad de información.

Se debe tener en cuenta que, aunque se tarda en construir los modelos, la clasificación de los datos es rápida, ya que sólo es necesario calcular los atributos BIF y la distancia de Mahalanobis (o el resultado SVM). Estos cálculos son rápidos y principalmente son operaciones entre píxeles (excepto el cálculo de los atributos básicos de imagen). Por tanto, las tareas pueden ser paralelizadas fácilmente e implementadas para funcionar rápidamente, por ejemplo, en una unidad de procesamiento gráfico (GPU). La rapidez es muy importante cuando se quiere implantar el control de calidad en una línea de producción.

SVM no funcionará adecuadamente si aparecen en la comida objetos extraños 'nuevos' (distintos de aquellos que han sido utilizados en el entrenamiento). Como no es posible crear un conjunto de entrenamiento perfecto, que represente todos los objetos extraños posibles, trabajos futuros deberían centrarse en el estudio de GMM.

El número de Gaussianas, D , se ha elegido mirando exclusivamente los valores de los atributos de intensidad, pero realmente se han añadido muchos más atributos, además de los correspondientes a la intensidad de las modalidades. Por ello, este parámetro debería ser optimizado dividiendo los datos de manera que haya más de un conjunto de datos para validación. Asimismo, se debería realizar una detección basada en objetos para una clasificación más robusta, además de para proporcionar una estimación de los tamaños de objetos que podrían ser detectados.

Adicionalmente, los modelos se podrían testear con datos 'nuevos', es decir, datos de comida que contengan huesos u otros objetos extraños que no han sido incluidos en el desarrollo de los modelos. De hecho, también se tomaron en Munich datos para tests de esta naturaleza, los cuales contienen cuerpos extraños que no han sido utilizados para el entrenamiento ni para la optimización de los parámetros de los modelos.

En este caso de prueba, Ridge no merece la pena. Si quisiéramos crear modelos más complejos, con mayor número de atributos espaciales, que captasen mejor las texturas complejas, como la del pan de centeno, entonces sí que se debería plantear la selección de atributos. Esto es debido a que, al añadir más escalas y parámetros de alisado al análisis de texturas, la dimensionalidad crece rápidamente. Entonces, una red elástica podría tratar de eliminar escalas y parámetros de alisado que no aportasen información. Así, se reduciría el tiempo de cálculo de los atributos BIF, y podría implementarse el sistema en una línea de producción.

Resumiendo, la interferometría basada en rejillas de difracción (GBI) permite detectar materia orgánica y proporciona resultados de detección mucho mejores que un sistema de detección basado en rayos-X convencionales cuando se pueden encontrar como objetos extraños tanto objetos orgánicos como no orgánicos. Como consecuencia, una solución de control de calidad implementada con GBI en una línea de producción sería mucho más eficiente que una solución basada en rayos-X convencionales, para industrias donde es posible encontrar objetos extraños orgánicos.

Bibliografía

- [Arivazhagan and Ganesan, 2003] Arivazhagan, S. and Ganesan, L. (2003). Texture classification using wavelet transform. *Pattern recognition letters*, 24(9-10):1513–1521.
- [Batchelor et al., 2004] Batchelor, B. G., Davies, E. R., and Graves, M. (2004). Using x-rays to detect foreign bodies. In Edwards, M., editor, *Detecting foreign objects in food*, pages 226–264. Woodhead Publishing Ltd.
- [Bech, 2009] Bech, M. (2009). *X-ray imaging with a grating interferometer*. PhD thesis, University of Copenhagen, Niels Bohr Institute.
- [Bose, 1993] Bose, S.-N. (1993). Mahalanobis, prasanta, chandra. *Current Science*, 65(1):96–97.
- [Cartensen, 1992] Cartensen, J. M. (1992). *Description and simulation of visual texture, chapter 1*. PhD thesis, Technical University of Denmark, Department of Informatics and Mathematical Modeling.
- [Crosier and Griffin, 2008] Crosier, M. and Griffin, L. (2008). Texture classification with a dictionary of basic image features. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–7.
- [Crosier and Griffin, 2010] Crosier, M. and Griffin, L. (2010). Using basic image features for texture classification. *International Journal of Computer Vision*, 88(3):447–460.
- [Dempster et al., 1977] Dempster, A., Laird, N., and Rubin, D. (1977). Maximum likelihood from incomplete data via em algorithm. *Journal of the Royal statistical society series b-methodological*, 39(1):1–38.

- [Hastie et al., 2008] Hastie, T., Tibshirani, R., and Friedman, J. (2008). Support vector machines and flexible discriminants. In *The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd edition*. Springer.
- [Hooke, 1665] Hooke, R. (1665). *Micrographia*. Royal Society.
- [Khan, 2008] Khan, S. (2008). Ethem alpaydin. introduction to machine learning (adaptive computation and machine learning series). *Natural Language Engineering*, 14.
- [Kottler et al., 2010] Kottler, C., Revol, V., Kaufmann, R., Urban, C., Knop, K., Sennhauser, U., Jerjen, I., Luthi, T., Cardot, F., and Niedermann, P. (2010). Phase sensitive x-ray imaging: Towards its interdisciplinary applications. *AIP conference proceedings*, 1236:213–218.
- [Leung and Malik, 1999] Leung, T. and Malik, J. (1999). Recognizing surfaces using three-dimensional textons. 2:1010–1017 vol.2.
- [McDonald, 2009] McDonald, G. C. (2009). Ridge regression. *Wiley Interdisciplinary Reviews: Computational Statistics*, 1(1):93–100.
- [Nielsen, 2012] Nielsen, M. S. (2012). Novel x-ray imaging modalities - seeing through food. Master’s thesis, University of Copenhagen, Niels Bohr Institute.
- [Nielsen et al., 2012] Nielsen, M. S., Lauridsen, T., Christensen, L. B., and Fidenhans, R. (2012). X-ray dark-field imaging for detection of foreign bodies in food. *Elsevier*.
- [Pfeiffer et al., 2006] Pfeiffer, F., Weitkamp, T., Bunk, O., and David, C. (2006). X-ray dark-field imaging for detection of foreign bodies in food. *Nature Physics*, 2.
- [Röntgen, 1895] Röntgen, W. C. (1895). Über eine neue art von strahlen. In *Sitzungsberichte der Physikalisch-Medizinischen Gesellschaft in Würzburg*, volume 137, pages 132–141.
- [Rost and Oldfield, 2000] Rost, F. W. D. and Oldfield, R. J. (2000). *Photography with a microscope*. Cambridge University Press.
- [Scherer et al., 2014] Scherer, K., Birnbacher, L., Chabior, M., Herzen, J., Mayr, D., Grandl, S., Sztrókay-Gaul, A., Hellerhoff, K., Bamberg, F., and Pfeiffer, F. (2014). Bi-directional x-ray phase-contrast mammography. *PloS One*, 9(5):e93502.
- [Spiegel, 1995] Spiegel, P. K. (1995). The first clinical x-ray made in america-100 years. In *AJR Am J Roentgenol*, volume 164, pages 241–243.

- [Takashi, 2009] Takashi, A. (2009). Detection for foreign bodies and bones in meat and meat products. In *International Congress of Meat Science and Technology*, 55, Copenhagen, Denmark.
- [Toledo, 2014] Toledo, M. (2014). X-ray safeline, www.mt.com/safeline-xray.
- [Zernike, 1942] Zernike, F. (1942). Phase contrast, a new method for the microscopic observation of transparent objects part ii. In *Physica*, volume 9, pages 974–986.