



Universidad Nacional Autónoma de México

POSGRADO EN CIENCIA E INGENIERÍA DE LA COMPUTACIÓN

**SEGMENTACIÓN AUTOMÁTICA DE LESIONES EN
IMÁGENES DE ULTRASONIDO DE MAMA
UTILIZANDO REDES NEURONALES
CONVOLUCIONALES**

T E S I S

Que para optar por el grado de:

Maestra en Ciencia e Ingeniería de la Computación

Presenta:

Vivian Bass Vega

Tutor Principal:

Dr. Jorge Alberto Márquez Flores

ICAT - UNAM

Ciudad Universitaria, Cd.Mx., Marzo 2022



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

*A mis padres, a mi hermano y a Dennis.
Gracias infinitas.*

Agradecimientos

En primera instancia, agradezco profundamente a mi asesor de tesis, al doctor Jorge A. Márquez, por todo el apoyo que me ha brindado en el tiempo de conocerlo, por sus palabras de aliento cuando todo parecía perdido y por ser mi guía durante el trabajo realizado. Gracias infinitas por su tiempo, comprensión, paciencia, consejos, amistad y conocimientos inyectados en esta tesis.

A los miembros del Comité Tutor: Dr. Boris Escalante, Dr. Alfonso Gastélum y Dr. Gibrán Fuentes, por su apoyo y comprensión en esta etapa, y por sus comentarios que han enriquecido el contenido de este trabajo.

Un especial agradecimiento al Dr. Iván M. Rosado, quien sin su guía y apoyo este trabajo no hubiera sido posible. Gracias por permitirme ser parte del grupo académico más talentoso que he conocido.

Gracias a mis padres, Conchita y Rodrigo, por darme su apoyo incondicional, por ser mis guías de vida y mi refugio, y en especial gracias por enseñarme a no darme por vencida. Los amo.

A mi hermano, Rodrigo, por no dejar de creer en mi y ser mi ejemplo a seguir. Gracias por siempre estar dispuesto a apoyarme y por las incontables risas juntos.

A Dennis, por ser mi cómplice, mi compañero de vida y mi luz al final del túnel. Gracias por tu apoyo, tus palabras, por ser mi motivación, por estar a mi lado cuando más lo he necesitado, por enseñarme a no preocuparme por las cosas que no puedo controlar y por sufrir este proceso conmigo. Sin duda no lo habría logrado sin ti.

A mis amigos, Fer, Daniel, Ponce, Monse, Jony y Julio, por estar conmigo desde el primer semestre de la Facultad y estar siempre para disfrutar, quejarnos, reír y llorar. Gracias por su apoyo incondicional y sus ánimos. Su amistad es invaluable.

Gracias especiales a Julieta Mateos, quien también ha sido mi guía y apoyo en este proceso, mi compañera de trabajo y mi maestra. Gracias por todo lo que me has enseñado, lo que nos hemos reído y por las pláticas siempre tan necesarias e interesantes.

A Noye, por hacerme reír, por enseñarme a ser una mejor persona y por darme el último empujón que necesitaba.

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca recibida durante la maestría con número de CVU 1003600.

Al Proyecto Ciencia de Frontera “Imágenes radiológicas cuantitativas para la caracterización no invasiva de cáncer de mama” número 1311307 financiado por CONACYT, y gracias al cual esta investigación fue posible.

Expreso mi profundo agradecimiento al Posgrado en Ciencia e Ingeniería de la Computación y a la Universidad Nacional Autónoma de México por todos los conocimientos adquiridos y el apoyo que me brindaron durante la maestría.

Resumen

El cáncer de mama es la principal causa de muerte por cáncer en mujeres en México y alrededor del mundo. Esto se debe principalmente al diagnóstico tardío y al precio del tratamiento del cáncer.

El ultrasonido es una de las herramientas basadas en imágenes más utilizadas para la evaluación de esta enfermedad, ya que puede ayudar a discriminar masas sólidas frente a quísticas, así como entre masas benignas o malignas. Sin embargo, analizar las imágenes por ultrasonido y producir diagnósticos acertados, depende en gran medida de la experiencia que tengan los radiólogos analizando imágenes, por lo que un análisis que sea independiente de la experiencia del radiólogo debería producir un mejor diagnóstico; por lo tanto, es necesario desarrollar sistemas de detección automática de lesiones para imágenes de ultrasonido. Estos sistemas se basan en la segmentación de imágenes, que es una técnica de procesamiento de imágenes que se utiliza para analizar y agrupar píxeles por sus características. La segmentación de imágenes de ultrasonido representa desafíos importantes debido a la complicada apariencia de tejido sano y tumoral en la imagen, que incluye patrón de moteado (conocido como ruido *speckle*, en inglés), bajo contraste, bordes borrosos, etc.

Para la realización de este trabajo se cuenta con una colección de imágenes formada por imágenes de señales de radiofrecuencia sin procesamiento, conocidas como imágenes de investigación, adquiridas a través de la licencia de investigación Axius Direct del equipo Siemens Acuson S2000 de pacientes del Instituto Nacional de Cancerología (INCan).

Para poder utilizar estas imágenes posteriormente, se necesita que las radiólogas del INCa realicen una segmentación manual de éstas. Como se explicará más adelante, para que las radiólogas puedan hacer una segmentación precisa y reproducible necesitan que las imágenes de investigación sean lo más parecido posible a las imágenes clínicas con las que ellas acostumbran trabajar, por lo tanto, para mejorar el aspecto de las imágenes de investigación, se propone utilizar un autocodificador, esto con el objetivo de disminuir el patrón de moteado de las imágenes, cuidando que se conserve en buen estado la estructura de la mama.

Posteriormente se realiza la réplica de la red neuronal STAN (Small Tumor Aware Network) [33] y se entrena con las imágenes de investigación, para poder comparar el desempeño de la red con la segmentación realizada por dos radiólogas del INCa. De esta forma se puede determinar si los resultados predichos por la red son parecidos a los que podría proporcionar otro radiólogo.

Los resultados obtenidos aplicando la arquitectura STAN a las imágenes de investigación procesadas con el autocodificador y realizando aumento de datos son: exactitud de 0.96, índice Dice de 0.91, sensibilidad de 0.93, especificidad de 0.99 y precisión de 0.99. Mientras que los resultados obtenidos aplicando la arquitectura STAN a las imágenes de investigación con las imágenes procesadas con el autocodificador, realizando aumento de datos e inicializando la red con los pesos y sesgos entrenados en el modelo anterior son: exactitud de 0.96, índice Dice de 0.91, sensibilidad de 0.94, especificidad de 0.98 y precisión de 0.98. Por otro lado, la evaluación entre las radiólogas indica: exactitud de 0.95, índice Dice de 0.92, sensibilidad de 0.92, especificidad de 0.98 y precisión de 0.98. Por lo anterior, se puede concluir que las predicciones de los modelos basados en redes neuronales evaluados llegan a ser muy parecidas a las segmentaciones que podría realizar un radiólogo.

Se espera que con ayuda de este trabajo se pueda desarrollar un sistema en línea de detección de lesiones en imágenes de ultrasonido de mama, que ayude a los radiólogos a hacer mejores diagnósticos.

Índice general

1. Introducción	1
1.1. Presentación	1
1.2. Planteamiento del problema	2
1.3. Objetivo general	3
1.4. Objetivos específicos	3
1.5. Contribución y relevancia	4
1.6. Metodología	4
1.7. Estructura del trabajo	5
2. Conceptos Generales	7
2.1. Cáncer de mama	7
2.2. Diagnóstico del cáncer de mama	8
2.2.1. Estudios de imagen	9
2.2.1.1. Mamografía	9
2.2.1.2. Ultrasonido	10
2.2.1.3. Resonancia magnética	14
2.2.2. Biopsia	15
2.3. Segmentación de imágenes	16
3. Estado del arte	17
3.1. Reducción de patrón de moteado en imágenes de ultrasonido	17
3.2. Segmentación por umbralización	18
3.3. Segmentación por agrupación	21
3.4. Segmentación por contornos activos	23
3.5. Segmentación con redes neuronales	26

4. Metodología	31
4.1. Adquisición de datos	31
4.2. Preparación de datos	32
4.3. Aplicación para la segmentación manual	33
4.4. Métodos propuestos	37
4.4.1. Reducción de patrón de moteado en imágenes de investigación	37
4.4.1.1. Autocodificadores	37
4.4.2. Segmentación de lesiones en imágenes de investigación	42
4.4.2.1. Redes neuronales convolucionales: Small Tumor Aware Network (STAN)	43
4.4.2.2. Aumento de datos	45
4.4.3. Validación de los métodos	46
5. Resultados	51
5.1. Máscara de segmentación	51
5.2. Reducción de patrón de moteado en imágenes de investigación	53
5.3. Segmentación de lesiones en imágenes de investigación	56
5.3.1. Arquitectura STAN con imágenes clínicas	56
5.3.2. Arquitectura STAN con imágenes de investigación con reducción de patrón de moteado	58
5.3.3. Arquitectura STAN con dos <i>datasets</i> realizando inicialización de los pesos	62
5.3.4. Cálculo de errores entre radiólogas	65
5.3.5. Comparación del desempeño de los distintos modelos con el error entre radiólogas	67
5.4. Discusión de resultados	67
6. Conclusiones	71
6.1. Trabajo futuro	73
A. Apéndice	75
A.1. Diagrama de aplicación para la segmentación manual	76
Bibliografía	77

Introducción

1.1. Presentación

El cáncer de mama es el cáncer más común en mujeres en todo el mundo y es la principal causa de muerte por cáncer. En 2019, hubo aproximadamente 1.7 millones de casos nuevos al año y 522,000 muertes relacionadas con esta enfermedad [9].

En México, el cáncer de mama ha aumentado de manera constante durante los últimos 30 años. Esto se debe a muchos factores, incluido el envejecimiento de la población, la falta de educación e información relacionada con la enfermedad, la falta de tratamiento oportuno, las demoras en el tratamiento en los hospitales públicos, falta de especialistas, escasez de insumos para el tratamiento, etc. [9]. Por este motivo, se están mejorando las tecnologías para la detección oportuna del cáncer, como la mamografía y los estudios por imagen de ultrasonido [22].

La mamografía es la técnica más utilizada y eficaz para detectar el cáncer de mama, pero debido a que las lesiones tienen la misma atenuación que el tejido denso, éstas pueden quedar ocultas [41], por lo que existen limitaciones para realizar mamografías en senos densos (Figura 1.1).

La densidad mamaria es un factor importante en la detección y el diagnóstico del cáncer de mama porque reduce su sensibilidad y muestra un riesgo significativamente mayor de desarrollar la patología (4.7 veces en comparación con las mujeres con mamas grasas) [9].

Actualmente, la alternativa a la mamografía es el ultrasonido, que se utiliza como

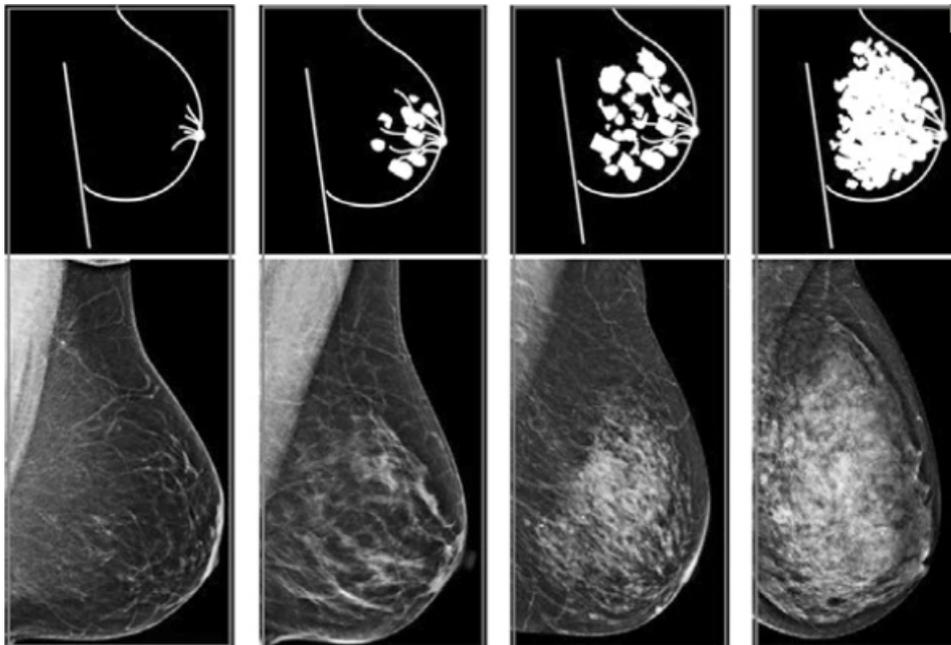


Figura 1.1: Las imágenes muestran la densidad de la mama en la mamografía [4].

complemento en la detección del cáncer de mama debido a su seguridad y sensibilidad [36]. El ultrasonido puede distinguir entre quistes y nódulos sólidos, y de éstos, los benignos de los malignos, pero esto depende en gran medida de la experiencia del radiólogo que realice la interpretación de las imágenes.

Para interpretar la imagen, el radiólogo debe detectar la lesión y luego segmentarla, es decir, separar la lesión del fondo. La segmentación de la lesión ayuda al radiólogo a saber si hay una masa en la mama y dónde está ubicada, así como saber si un tumor es potencialmente maligno y determinar si el paciente necesita una biopsia. La segmentación es el primer paso para planificar el tratamiento así como saber si un tratamiento está funcionando al observar la evolución del tumor.

1.2. Planteamiento del problema

Un número significativo de radiólogos continúa realizando la segmentación de imágenes mediante apreciación visual. Lo anterior hace que su trabajo sea tedioso, lento y dependiente de la experiencia del usuario.

Una interpretación que no dependa de la experiencia del radiólogo reduce la probabilidad de que los médicos malinterpreten la malignidad de la lesión por lo que existe la necesidad de desarrollar un sistema de segmentación de lesiones que apoye al radiólogo a la identificación de las regiones de interés.

Los sistemas de detección automáticos y semiautomáticos están basados en la segmentación de imágenes, que es una técnica de procesamiento que se utiliza para analizar y agrupar píxeles según características comunes. La segmentación de las imágenes de ultrasonido representa un problema desafiante debido a la naturaleza de la imagen, que incluye patrón de moteado, bajo contraste, bordes borrosos, sombras, etc.

Debido a lo anterior, se han probado varios métodos de segmentación de imágenes de ultrasonido que permiten desarrollar sistemas de diagnóstico asistidos por computadora que ayuden a los radiólogos a interpretar mejor las imágenes.

1.3. Objetivo general

- Realizar un sistema de detección automático de lesiones en imágenes de investigación de ultrasonido de mama que pueda ser utilizado para segmentar las lesiones en las imágenes del proyecto “Imágenes radiológicas cuantitativas para la caracterización no invasiva de cáncer de mama” número 1311307 financiado por CONACYT.

1.4. Objetivos específicos

- Mejorar una aplicación desarrollada en Swift¹ para realizar el marcado de las imágenes por parte de las radiólogas.
- Diseñar e implementar un algoritmo que permita filtrar el patrón de moteado para tener imágenes más claras de las lesiones de mama.
- Obtener los contornos de referencia de las imágenes de ultrasonido realizado por radiólogas del Instituto Nacional de Cancerología (INCan).

¹Lenguaje de programación creado por Apple enfocado en el desarrollo de aplicaciones para iOS y macOS.

- Diseñar e implementar un algoritmo que permita obtener de forma automática las máscaras que se obtendrán de los contornos de referencia.
- Replicar el enfoque propuesto en “STAN: Small Tumor Aware Network” [33] con las imágenes proporcionadas por el INCan.
- Validar el método de segmentación.

1.5. Contribución y relevancia

Este trabajo tiene como objetivo desarrollar un algoritmo que ayude a filtrar el patrón de moteado en las imágenes de ultrasonido a través de un autocodificador, los cuales son redes neuronales con dos componentes: un codificador que recibe la entrada y la comprime y un decodificador que toma la versión comprimida de los datos y la reconstruye, de manera que al tener control total sobre los valores de entrada y los valores de salida es posible entrenar el modelo de manera que se elimine en gran medida el ruido y de esta forma lograr que las radiólogas con las que trabajamos puedan utilizar esas imágenes para segmentarlas manualmente.

Por otro lado, con las imágenes segmentadas manualmente se busca construir una colección de imágenes para entrenar un modelo de segmentación automática de imágenes de ultrasonido de mama mediante una red neuronal convolucional, la cual deberá ser capaz de detectar lesiones de cualquier tamaño en imágenes de investigación.

Esa segmentación automática se validará con las imágenes marcadas por las radiólogas que fueron apartadas del conjunto inicial, midiendo el error entre la segmentación hecha por las radiólogas contra el método propuesto. De igual manera se calcularán la sensibilidad, especificidad, precisión, exactitud y criterios de similitud para saber cuánto se parece el área segmentada por la red con el área marcada por las expertas.

1.6. Metodología

Se investigaron distintos métodos de segmentación utilizados en imágenes de ultrasonido. Una vez que se revisó el estado del arte, se decidió utilizar redes neuronales convolucionales para realizar la segmentación automática de lesiones en imágenes de ultrasonido de mama.

Debido a que ya se contaba anteriormente con las imágenes adquiridas en el INCAn, se trabajó en una herramienta para que las radiólogas pudieran realizar la segmentación manual de las lesiones de una forma sencilla para poder utilizar esos datos como las imágenes de entrenamiento y validación para los modelos. Además, se desarrolló un programa para poder obtener las máscaras de segmentación a partir la imagen obtenida con la herramienta anteriormente mencionada.

Ya que en la clínica se trabaja con imágenes más nítidas que las imágenes de investigación con las que se cuenta, fue necesario reducir el patrón de moteado de éstas para poder realizar la segmentación manual con imágenes que fueran lo más parecido posible a las imágenes clínicas.

Posteriormente se replicó el enfoque propuesto en STAN [33] y se evaluaron los modelos entrenados con dos conjuntos de datos distintos.

Finalmente se validaron ambos métodos de segmentación y se obtuvo el error de segmentación entre radiólogos para comparar los resultados.

1.7. Estructura del trabajo

En el capítulo 2 se describen los conceptos generales para el desarrollo de este trabajo. Aquí se presenta la importancia del diagnóstico de cáncer de mama, por qué el ultrasonido es la mejor alternativa a la mamografía, cómo surge el patrón de moteado en las imágenes de ultrasonido, por qué es importante reducir el patrón de moteado en este trabajo y qué es la segmentación de imágenes en la medicina.

En el capítulo 3 se presenta el estado del arte de la reducción de patrón de moteado en imágenes de ultrasonido, así como las técnicas que se han utilizado para la segmentación en este tipo de imágenes. Cada técnica se describe y se presentan sus ventajas y desventajas.

En el capítulo 4 se presenta el desarrollo y mejora de una aplicación para iPad para realizar el marcado de las imágenes por parte de las radiólogas, el método de adquisición de datos, el pre-procesamiento de las imágenes utilizadas, y se describen los métodos propuestos para la reducción de patrón de moteado, así como la segmentación

1. INTRODUCCIÓN

de lesiones en imágenes de ultrasonido.

En el capítulo 5 se presentan los resultados obtenidos con el método propuesto para la reducción de patrón de moteado, tres experimentos diferentes con la arquitectura utilizada para la segmentación automática de imágenes, las métricas utilizadas para calcular el error entre radiólogas y finalmente la comparación del desempeño de la red utilizada con el error de segmentación entre radiólogas.

En el capítulo 6 se analizan los resultados obtenidos en el capítulo 5 al implementar los métodos propuestos. Finalmente, se presentan opciones para realizar trabajo futuro en la misma dirección que esta tesis.

Conceptos Generales

En este capítulo se revisarán las bases para el desarrollo de esta tesis. Primero se profundizará en el cáncer de mama y la forma de diagnosticarlo, así como los distintos estudios por imagen que se pueden realizar para detectarlo. Además, se profundizará en lo que es el ultrasonido y cómo se genera, se da una explicación de lo que es el patrón de moteado y finalmente se explica qué es y cuál es el objetivo de la segmentación de imágenes.

2.1. Cáncer de mama

El cáncer de mama es la segunda enfermedad más común en el mundo después del cáncer de pulmón y la principal causa de muerte por cáncer en mujeres. La mortalidad por cáncer de mama sigue siendo un problema de salud pública mundial, principalmente en países con recursos limitados [13], a pesar de que las pruebas de diagnóstico y terapéuticas han mejorado en los últimos años. Gracias al diagnóstico y tratamiento tempranos en países con ingresos altos, la supervivencia del cáncer de mama durante los primeros 5 años después del diagnóstico es mayor al 90 %, mientras que en países con ingresos bajos la tasa de supervivencia solamente es entre el 40 % y el 66 % [3].

Para 2020, 2,3 millones de mujeres fueron diagnosticadas con cáncer de mama en todo el mundo y se estima que 685,000 fallecieron por esta enfermedad. El cáncer de mama afecta a mujeres de todas las edades después de la pubertad, pero las probabilidades de contraer cáncer de mama aumentan conforme pasan los años. El cáncer de mama no es una enfermedad contagiosa o infecciosa, a diferencia de algunos cánceres como cáncer de cuello uterino y cánceres cuya etiología está asociada con el proceso de

infección, como la infección por el virus del papiloma humano (VPH). No se conocen infecciones virales o bacterianas asociadas con el desarrollo del cáncer de mama.

Aproximadamente la mitad de los casos de cáncer de mama son mujeres sin factores de riesgo identificables, excepto el sexo (mujer) y la edad (40+). Ciertos factores aumentan el riesgo de desarrollar la enfermedad, incluido el envejecimiento, la obesidad, el uso excesivo de alcohol, los antecedentes familiares, los antecedentes de exposición a la radiación, los antecedentes reproductivos, el tabaquismo y la terapia hormonal posmenopáusica [3].

Entre 1930 y 1970, la mortalidad por cáncer de mama cambió ligeramente. La mejora de la supervivencia comenzó en la década de 1980 en países con programas de detección temprana combinados con varios tipos de tratamientos para controlar la enfermedad.

2.2. Diagnóstico del cáncer de mama

El síntoma clínico más común es la presencia de un bulto normalmente indoloro descubierto por el propio paciente incidentalmente, o por pruebas de laboratorio de rutina. Puede asociarse con otros signos clínicos como cambio de tamaño del seno, forma, apariencia, enrojecimiento, agrietamiento, cambio en la forma del pezón y la piel circundante, secreción inusual, etc. Los cánceres en etapa tardía pueden erosionar la piel y causar llagas abiertas, que no siempre son dolorosas.

El cáncer de mama puede extenderse a otras partes del cuerpo y causar otros síntomas. A menudo, el primer sitio de diseminación más común son los ganglios linfáticos axilares, pero puede haber ganglios linfáticos cancerosos que no se detectan.

El cáncer de mama puede manifestarse de muchas formas, por lo que es aconsejable un examen físico completo. Las mujeres con anomalías persistentes deben hacerse una mamografía y, en algunos casos, algunas pruebas, como una biopsia, para determinar si el bulto es maligno o benigno. Este diagnóstico se basa en la evaluación clínica (antecedentes personales y familiares), pruebas con imágenes y biopsia.

2.2.1. Estudios de imagen

Las pruebas de imagen como la mamografía, ultrasonido y la resonancia magnética permiten la detección, caracterización, evaluación de la enfermedad y seguimiento de las lesiones mamarias.

2.2.1.1. Mamografía

La mamografía es la técnica más utilizada para diagnosticar lesiones sospechosas. Se puede dividir en dos categorías: escrutinio y diagnóstico. Las mamografías de escrutinio son estudios de mujeres sanas para detectar los primeros signos de cáncer de mama. La mamografía de diagnóstico se usa para detectar el cáncer después de que se encuentra un bulto u otro signo o síntoma de cáncer. Las pruebas de diagnóstico brindan información sobre la ubicación, el número y las características de las lesiones palpables [31].

Algunas de las anomalías que los radiólogos buscan al analizar imágenes de mamografías son calcificaciones, masas, asimetría en la densidad o distorsiones estructurales. Dependiendo de las características de las anomalías encontradas, el radiólogo brindará su interpretación y una serie de recomendaciones.

Las mamas están formadas por tejidos muy similares, y las lesiones que los radiólogos buscan como indicio de la posibilidad de un tumor son muy pequeñas o similares al tejido normal, por lo que la obtención de una imagen radiológica de la mama es un desafío [8]. Además, una de las principales desventajas de la mamografía es que los senos densos aumentan las posibilidades de que se pase por alto el cáncer de mama ya que el tejido denso puede enmascarar a las lesiones (Figura 2.1); y aunque se desconoce la razón exacta, los senos densos aumentan el riesgo de cáncer de mama.

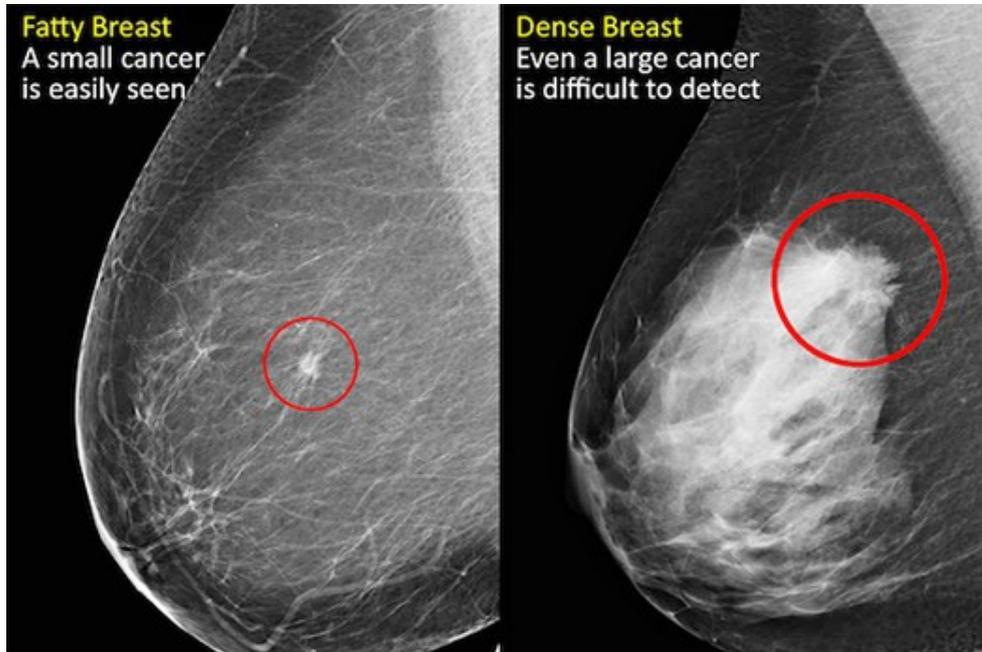


Figura 2.1: Izq: La lesión se puede ver muy claramente. Der: La lesión no se ve aunque sea grande debido a la densidad de la mama [38].

2.2.1.2. Ultrasonido

El ultrasonido (US) es el método de detección complementario más utilizado para mujeres con senos densos (Figura 2.2) [38]. Es portátil, libre de riesgo de radiación, toma imágenes en tiempo real y es relativamente económico en comparación con otros métodos de obtención de imágenes como la resonancia magnética.

Debido a que en esta tesis se trabajarán con imágenes de ultrasonido, se profundizará más en este tema.

El ultrasonido tiene frecuencias por encima del límite de la audición humana, es decir, por encima de 20 kHz, y se realiza mediante el enfoque de pulso-eco [10]. Para obtener una imagen de ultrasonido, se debe enviar un pequeño pulso de ultrasonido desde el transductor al cuerpo. A medida que las ondas de ultrasonido atraviesan el tejido, algunas se transmiten a estructuras más profundas, otras se reflejan como ecos en el transductor, otras se dispersan y otras se convierten en calor.

La cantidad de eco que regresa después de tocar el tejido está determinada por

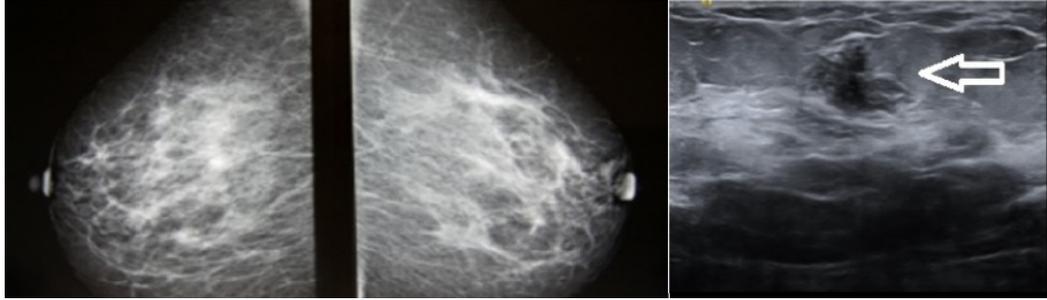


Figura 2.2: Izq: La lesión no se aprecia en mamografía.
Der: La lesión se detecta en US [2].

una propiedad llamada impedancia acústica. Esta es una propiedad física intrínseca del medio que se define como la densidad del medio multiplicada por la velocidad de propagación de la onda de ultrasonido en ese medio. La impedancia acústica de los órganos que contienen aire es la más baja y la impedancia acústica de órganos densos como los huesos es muy alta. La intensidad del eco reflejado es proporcional a la diferencia de impedancia acústica entre los dos medios; no se genera eco si las impedancias acústicas de los dos tejidos son iguales. Las interfaces de tejidos blandos con impedancia acústica similar suelen producir ecos de baja intensidad. Por el contrario, la interfaz entre el tejido blando y el hueso o el pulmón produce ecos muy fuertes debido a las diferencias en la impedancia acústica.

Cuando las ondas ultrasónicas entran en tejido corporal con diferentes impedancias acústicas, algunas se reflejan en el transductor (señal de eco), mientras que otras continúan entrando. Los ecos devueltos por varios pulsos consecutivos se procesan y combinan para producir una imagen. Por lo tanto, el transductor ultrasónico actúa como altavoz (produciendo ondas sonoras) y como micrófono (recibiendo ondas sonoras). El ultrasonido debe emitir pulsos con tiempo suficiente entre ellos para que la señal alcance el objetivo de interés y sea reflejada por el transductor como un eco antes de que se produzca el siguiente pulso.

El transductor ultrasónico contiene una gran cantidad de cristales piezoeléctricos que oscilan en respuesta a la corriente aplicada. Las ondas sonoras se pueden representar en términos de frecuencia, longitud de onda y amplitud. Las de alta frecuencia se utilizan para crear imágenes del tejido mamario. La imagen se crea cuando pasa la onda y es reflejada por el tejido (Figura 2.3).

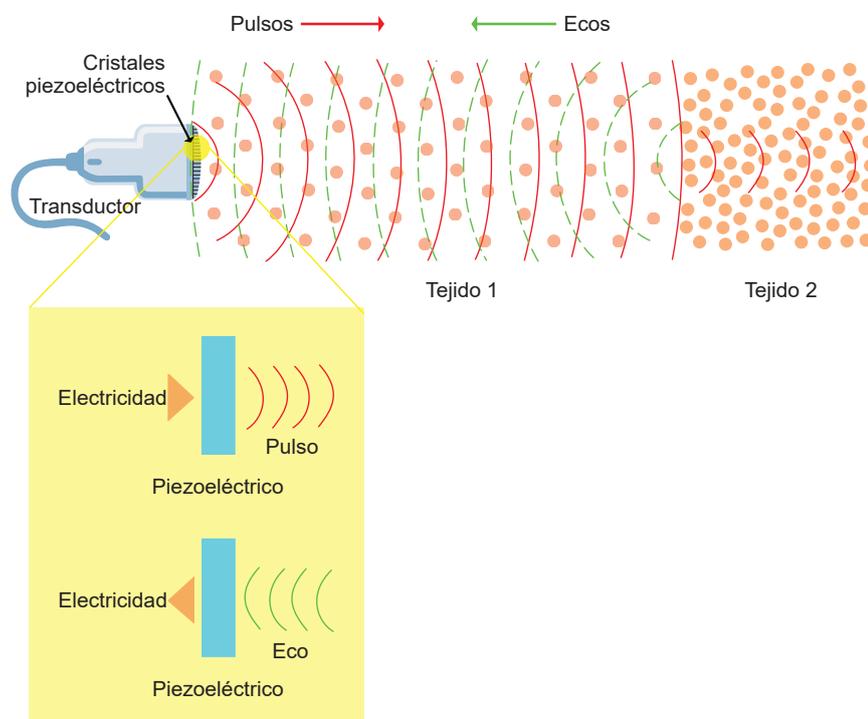


Figura 2.3: Diagrama del ultrasonido.

La longitud de onda y la frecuencia de ultrasonido están correlacionadas inversamente: el ultrasonido de baja frecuencia tiene una longitud de onda larga y viceversa. Las frecuencias altas producen imágenes de alta resolución y están más atenuadas que las frecuencias bajas a ciertas distancias, lo que las hace más adecuadas para obtener las estructuras superficiales. Por el contrario, las bajas frecuencias proporcionan imágenes de baja resolución, pero con menos atenuación, por lo que pueden penetrar estructuras más profundas.

A medida que el pulso ultrasónico atraviesa el tejido, su intensidad se debilita. Esta atenuación es el resultado de la reflexión y la dispersión. La atenuación también varía entre los tejidos corporales, con el grado más alto en los huesos, menor en los músculos y órganos sólidos, y el más bajo en la sangre para cualquier frecuencia dada. Todos los dispositivos de ultrasonido compensan intrínsecamente la atenuación promedio esperada aumentando automáticamente la ganancia en áreas más profundas de la pantalla (la intensidad de la señal que hace que la imagen se vea más brillante). Esta es la causa de un artefacto muy común conocido como "realce acústico posterior" que describe un

área relativamente hiperecoica posterior a los grandes vasos sanguíneos o quistes. Las estructuras que contienen fluidos tienen menos atenuación del sonido que las estructuras sólidas, por lo que la intensidad de un pulso después de pasar a través de un fluido es mayor que después de pasar a través de una cantidad equivalente de tejido sólido.

Patrón de moteado

Una característica inherente de la imagen de ultrasonido es la presencia de patrón de moteado. El moteado es un patrón de interferencia determinista aleatorio que degrada los bordes y los detalles de una imagen (Figura 2.4). Este fenómeno complica la detección de pequeñas lesiones y de bajo contraste. También reduce la precisión de las tareas de procesamiento de imágenes por ultrasonido, como la extracción de características, la segmentación, el registro y la clasificación [40].

La mayoría de los tejidos biológicos aparecen en las imágenes de ultrasonido como si estuvieran llenos de pequeñas estructuras. Esto se debe a que cuando un pulso atraviesa un tejido que contiene una gran cantidad de discontinuidades, se producen ecos en diferentes direcciones. Los ecos se superponen y provocan interferencias constructivas y destructivas para generar una señal. [30].

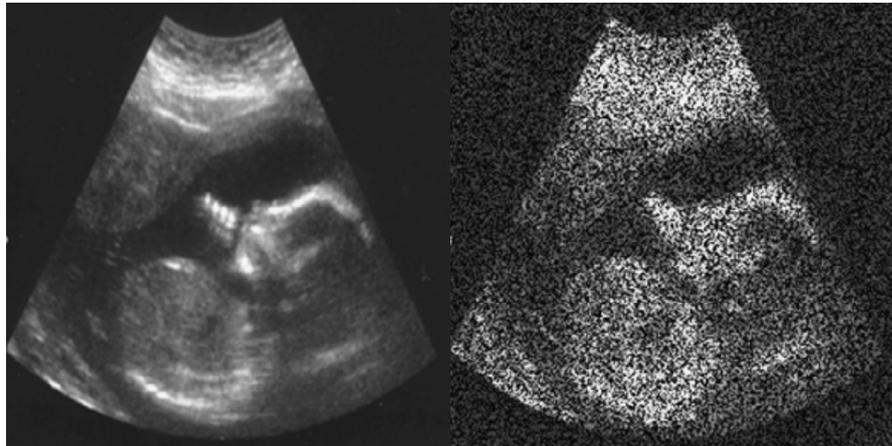


Figura 2.4: Izq: Ultrasonido fetal
Der: Ultrasonido fetal con patrón de moteado [23].

En este trabajo es importante la reducción del patrón de moteado debido a que las radiólogas con las que se trabajó tienen más experiencia realizando segmentación en imágenes clínicas, es decir, imágenes de ultrasonido que no tienen ruido. Como

2. CONCEPTOS GENERALES

se explicará en el capítulo 4, la colección de imágenes con la que se cuenta para la realización de esta tesis está formada por imágenes de señales de radiofrecuencia sin procesamiento, por lo que estas imágenes contienen mayor patrón de moteado que las imágenes clínicas.

Como se ha mencionado, una buena segmentación depende de la experiencia que tenga el radiólogo que analiza la imagen, por lo que se busca que al recabar las máscaras manuales, las radiólogas trabajen con imágenes parecidas a las que analizan a diario para no cometer errores en el marcado de las lesiones.

2.2.1.3. Resonancia magnética

Ésta es una técnica que utiliza imanes y ondas de radio para crear imágenes detalladas del tejido mamario. Generalmente se utiliza una solución de contraste inyectada en el brazo ya que el cáncer absorbe más contraste que el tejido normal circundante (Figura 2.5).

Para las mujeres con alto riesgo de cáncer de mama, la detección por resonancia magnética debe realizarse al mismo tiempo que la mamografía anual. La resonancia magnética no se recomienda como prueba de detección de cáncer de mama debido a que es más probable encontrar un cáncer que resulta no canceroso (falso positivo) que en una mamografía. En última instancia, esto puede hacer que las mujeres se sometan a pruebas y biopsias innecesarias [12].

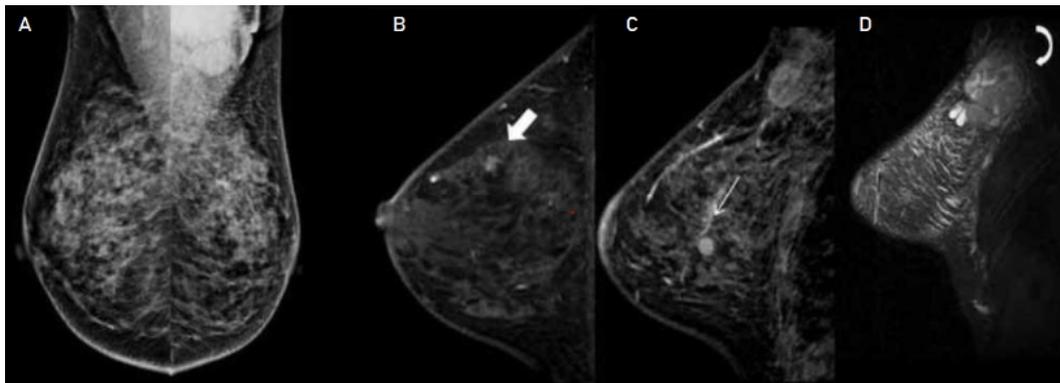


Figura 2.5: A: Mamografía sin lesiones. B, C, D: Resonancia que detecta nódulos irregulares [27].

2.2.2. Biopsia

Una biopsia de mama es un procedimiento en el que se obtiene una muestra de tejido mamario para conocer si existen células cancerígenas en el bulto que se identificó en la mama. Con los resultados de la biopsia se puede saber si un paciente necesita más tratamiento o algún procedimiento adicional [1].

Existen distintas formas de realizar una de biopsia de mama:

- Biopsia por succión con aguja fina: se utiliza para evaluar los bultos palpables durante un examen clínico de la mama. La succión con aguja fina es una forma rápida de distinguir entre quistes llenos de líquido y masas sólidas y evitar procedimientos de biopsia más invasivos; si la masa es sólida, se tomará una muestra de ese tejido.
- Biopsia con aguja gruesa: se utiliza para evaluar masas mamarias que se pueden ver en mamografías o ultrasonido, o masas que un médico puede sentir durante un examen clínico de las mamas. Un radiólogo o cirujano usa una aguja hueca delgada para tomar una muestra de tejido de tumor en la mama; en la mayoría de los casos, se realiza bajo control por ultrasonido.
- Biopsia estereotáctica: un radiólogo localiza la masa en la mama guiándose con mamografías, posteriormente inserta una aguja y toma muestras del tejido central de la lesión.
- Biopsia con aguja gruesa guiada por ultrasonido: un radiólogo localiza la masa en la mama guiándose con ultrasonido, posteriormente inserta una aguja y toma muestras del tejido central de la lesión.
- Biopsia con aguja gruesa guiada por resonancia magnética: un radiólogo localiza la masa en la mama guiándose con imágenes de resonancia magnética, posteriormente inserta una aguja y toma muestras del tejido central de la lesión.

El cáncer de mama es una enfermedad para la que los tratamientos están en constante evolución, mejorando el pronóstico de las pacientes por medio de la colaboración de todas las ramas de la oncología trazando distintos escenarios para obtener el mejor resultado posible.

En México, el sistema de salud es irregular y muchos centros carecen de los recursos necesarios para una atención moderna. Por ello, se incentiva la creación de equipos interdisciplinarios en todas las instituciones públicas y privadas con el fin de optimizar los recursos y obtener los mejores resultados de tratamiento en función de los factores de enfermedad, pacientes, tratamiento e institución.

2.3. Segmentación de imágenes

La segmentación de imágenes es una de las partes más importantes del procesamiento y análisis de imágenes, ya que busca separar regiones de interés (ROI) en función de características o atributos específicos. [11].

El objetivo de la segmentación es subdividir los objetos en la imagen; para la segmentación en imágenes médicas, los objetivos son:

- Estudiar estructuras anatómicas
- Identificar regiones de interés
- Realizar mediciones volumétricas de tejido para conocer el crecimiento o disminución tumoral
- Ayudar en la planificación del tratamiento

En los últimos 20 años se han propuesto muchas alternativas de segmentación automática y semiautomática para tumores mamarios, pero se ha detectado que no pueden segmentar masas o tumores muy pequeños, y detectar tumores pequeños es importante para encontrar cánceres en etapa temprana.

Existen muchos métodos de segmentación que son efectivos en determinados tipos de imagen. En el caso de las imágenes de ultrasonido, los métodos tradicionales a veces no funcionan de manera eficiente ya que éstas tienen patrón de moteado o ruido *speckle*, por el cual no es posible tener imágenes de buena calidad y no se pueden tener detalles tan finos, lo cual complica la segmentación automática de lesiones.

Estado del arte

En este capítulo se describen distintas técnicas que se han utilizado para la reducción del patrón de moteado en imágenes de ultrasonido, así como distintos métodos de segmentación de lesiones describiendo las ventajas y desventajas de su uso en imágenes de ultrasonido.

3.1. Reducción de patrón de moteado en imágenes de ultrasonido

En las imágenes de ultrasonido la intensidad del patrón de moteado puede ser alta, lo que provoca que se bloquee información relevante en la imagen, por lo que es importante la reducción de ese patrón. Ese filtrado se puede realizar tanto en el dominio espacial como en el de frecuencia. El objetivo de la técnica de reducción de moteado es eliminar el patrón de moteado de la imagen de tal forma que los bordes, la textura y otras características se conserven en la imagen, es decir, es importante equilibrar la cancelación de ruido con la capacidad del filtro para retener los detalles.

Los filtros lineales suavizan los bordes al reducir las manchas, por lo que no se recomiendan para reducir el moteado.

Los filtros Lee [26], Lee mejorado, Kuan [25], Frost [14], Gamma MAP [6], etc., son filtros de eliminación de ruido que se usan para limpiar imágenes de ultrasonido, pero pueden ser problemáticos si se necesita limpiar y conservar los bordes al mismo tiempo. Estos filtros están en el dominio espacial, se basan en estadísticas locales y se pueden considerar filtros paso bajas basados en el enfoque de error cuadrático medio,

que minimiza las motas a expensas de bordes suavizados.

Por otro lado, el filtro de vecino más cercano simétrico (Symmetric Nearest Neighbour: SNN) [15] explora el vecino más cercano y la simetría espacial para reducir la componente de moteado de la imagen. Del mismo modo, el filtro de difusión anisotrópico propuesto por Perona y Malik [42] es un esquema para detectar bordes y suavizar la imagen en múltiples escalas con el fin de mejorar el contraste y al mismo tiempo reducir el ruido. El filtro propuesto por Tomasi y Manduchi [39] es una alternativa al filtro de difusión anisotrópico, que es una combinación de dos núcleos gaussianos. Los detectores de bordes Canny se utilizan junto con los procesos de filtrado para mantener los bordes intactos, por ejemplo, se introdujo el filtro de difusión anisotrópica reductora de moteado (Speckle Reducing Anisotropic Diffusion: SRAD) combinado con un detector de bordes Canny, que mostró mejores resultados en la eliminación de motas [29].

3.2. Segmentación por umbralización

La umbralización es el método de segmentación más básico y sencillo, pero funciona muy bien con ciertos tipos de imágenes. Consiste en tomar un valor de intensidad de la imagen y establecerlo como un umbral basado en el histograma de la imagen para que éste sea un mínimo local del histograma. Los objetos en la imagen se separarán de acuerdo con la intensidad, por ejemplo: si la imagen tiene una resolución de 8 bits y el valor del píxel es mayor que el umbral, se le asignará un valor de 255 a ese píxel (el objeto), de otra manera, será cero (el fondo), convirtiendo la imagen original a una imagen binarizada.

Este método se puede dividir en umbralización global y umbralización local. El procesamiento por umbral global divide la imagen por un único umbral, como en el caso del algoritmo de Otsu. Con este algoritmo, al obtener el histograma de la imagen, la varianza de cada clase (objeto y fondo) es lo más pequeña posible, y al mismo tiempo lo más grande posible entre clases. Este método es útil si se tiene una iluminación uniforme en toda la imagen y un histograma bimodal, es decir, si ambos picos son claramente visibles (Figura 3.1).

Por otro lado, en el procesamiento de umbralización local, se eligen múltiples um-

brales porque hay un histograma multimodal (Figura 3.2) debido a cambios en la iluminación de la imagen, por lo que se divide la imagen en varias imágenes para encontrar el umbral apropiado para cada una, de modo que los objetos y el fondo se encuentren en función del umbral seleccionado para ese conjunto de píxeles.

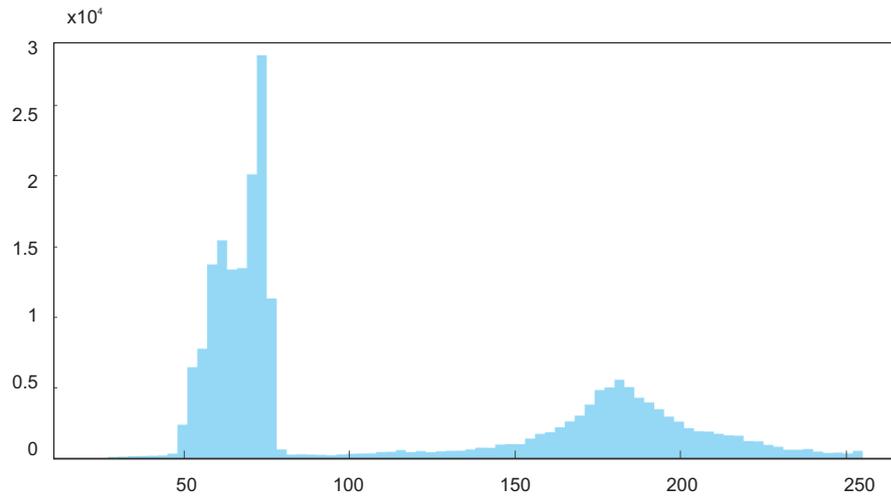


Figura 3.1: Ejemplo de histograma bimodal.

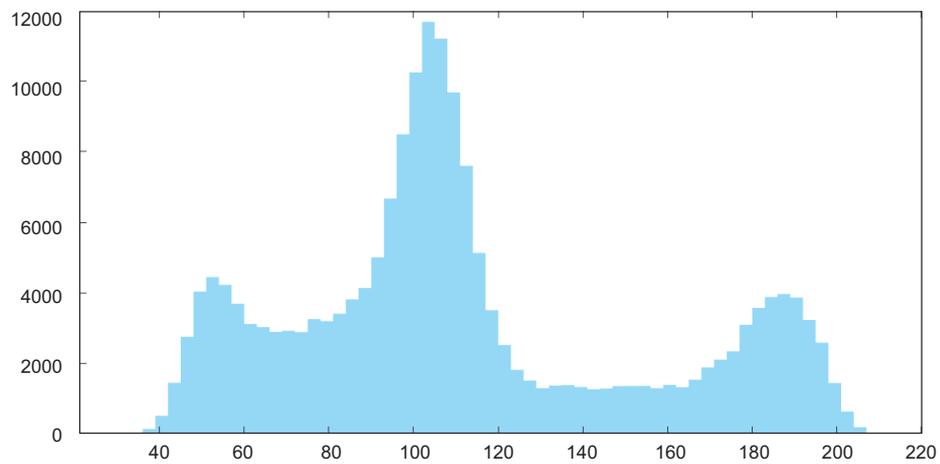


Figura 3.2: Ejemplo de histograma multimodal.

Este método no tiene en cuenta la posición de los píxeles, sólo los valores de intensidad. Este método funciona bien para imágenes con áreas claramente distinguibles y, como se describió anteriormente, donde hay una iluminación uniforme en toda la imagen (Figura 3.3).

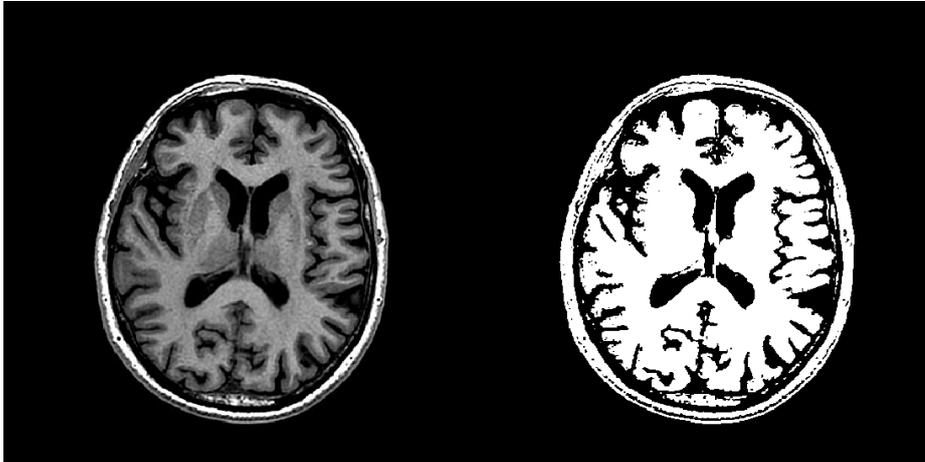


Figura 3.3: Segmentación de imagen de resonancia magnética.

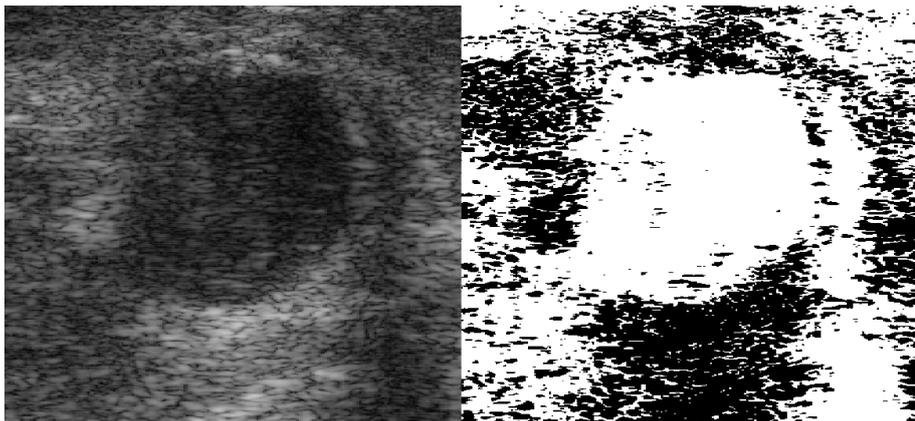


Figura 3.4: Segmentación en imagen de ultrasonido.

Para las imágenes de ultrasonido, se han realizado varias pruebas utilizando el método de umbralización, pero debido al patrón de moteado y que las intensidades de gris son muy similares, es muy difícil determinar el umbral correcto, ya que se puede eliminar información de la región de interés o bien, dejar mucha información que no es relevante para el caso de estudio (Figura 3.4).

3.3. Segmentación por agrupación

Este método intenta dividir el conjunto de datos (en este caso, la imagen) en n grupos según las características que comparten los píxeles. Esto hace que los píxeles de un grupo sean más similares a los píxeles de ese grupo que a los de otros grupos. A diferencia del método de umbralización, este método considera la posición al formar una región, es decir, además de que los píxeles sean similares, necesitan estar cerca unos de otros.

La segmentación por agrupación se puede dividir en dos subgrupos:

- **Agrupación dura:** en este tipo de agrupación, los píxeles pertenecen o no pertenecen a una región. Un algoritmo de este tipo de método es *K-means*. En imágenes médicas, este algoritmo se utiliza ampliamente para la segmentación y clasificación de imágenes. Éste es un algoritmo iterativo que tiene como objetivo dividir los píxeles de una imagen en k grupos, en los que cada píxel pertenece al grupo con el centroide más cercano minimizando la distancia euclidiana entre el píxel y el centroide del grupo.

El algoritmo inicia cuando se especifica el número de grupos k en los que se va a dividir la imagen (Figuras 3.5 y 3.6), se eligen al azar k píxeles como centroides iniciales, se calcula la distancia euclidiana de cada píxel a cada centroide y se asignará al grupo que minimice la distancia. Cuando se han asignado todos los píxeles a un grupo, se calcula el centroide de cada grupo y se repiten los pasos anteriores hasta que los píxeles ya no cambien de grupo. Cuando esto sucede, la segmentación de la imagen se completa.

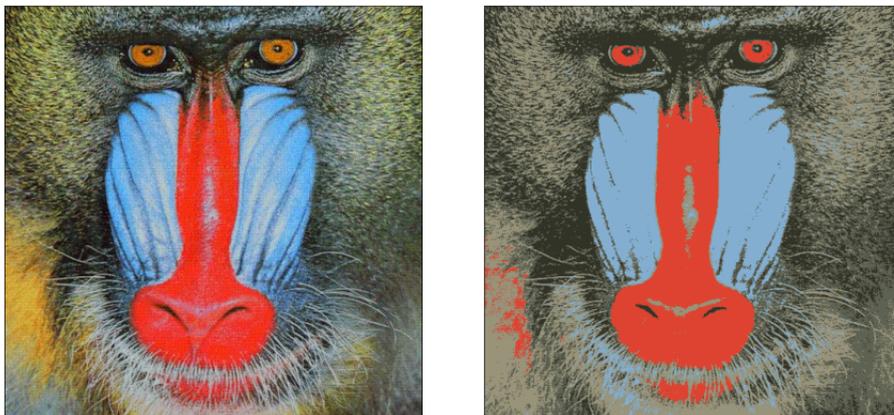


Figura 3.5: Segmentación con *k-means*, $k=3$.

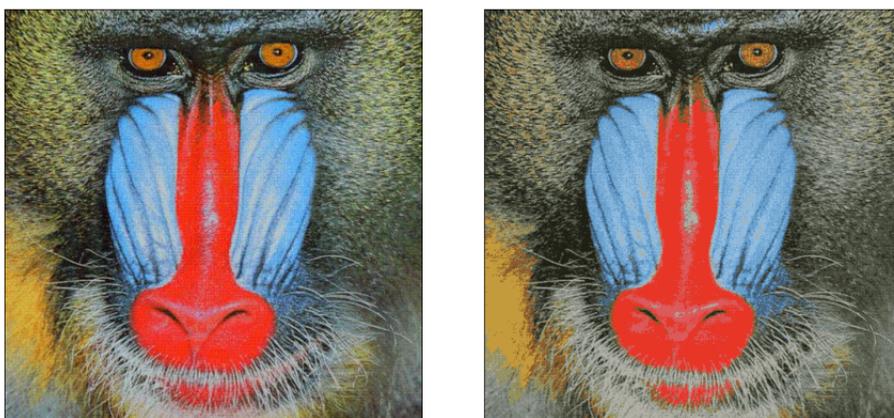


Figura 3.6: Segmentación con *k-means*, $k=15$.

- Agrupación suave: En este tipo de agrupación, a los píxeles se les asigna una probabilidad de pertenecer a cada grupo. Por lo general, cuando se utilizan métodos duros, es difícil clasificar los píxeles en un grupo u otro porque los píxeles se encuentran a una distancia similar de dos centroides y pueden tener atributos que pertenecen a diferentes grupos. El algoritmo *Fuzzy c-means* (FCM) se desarrolló para resolver este problema (Figura 3.7).

Este algoritmo asigna a cada píxel un valor de pertenencia para cada grupo, de modo que el mismo píxel pertenece a más de un grupo en cierto grado:

Primero, se debe seleccionar el número de grupos c para dividir la imagen. Luego, los centroides se asignan al azar, posteriormente se calcula el grado de pertenencia de los píxeles a los grupos, luego se vuelve a calcular el centroide para cada grupo y se calcula nuevamente el grado de pertenencia de los píxeles a cada grupo. Estos dos últimos pasos se repiten hasta que se alcanza el número máximo de iteraciones o hasta que los grupos ya no presenten cambios.

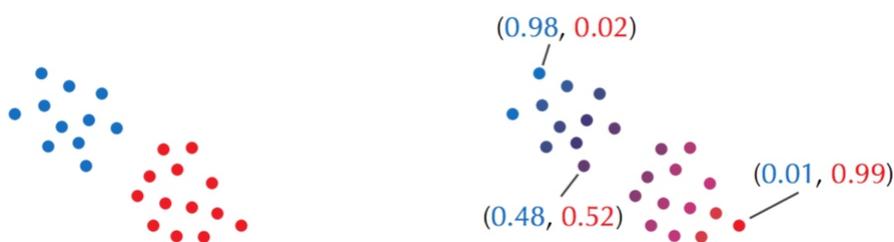


Figura 3.7: Diferencia entre segmentación con k-medias y FCM [17].

Para imágenes de ultrasonido, en [32] desarrollaron una nueva técnica de agrupamiento en la que se selecciona manualmente la región de interés. Esto sirve para que se analice solamente lo que está dentro de esa área y elimina la mayor parte del patrón de moteado. Luego, para reducir el moteado, crearon un algoritmo que filtra la imagen en el dominio de la frecuencia y calcula la acumulación de fase en la dirección con máxima energía. Luego utilizaron una variación del método *k-means* que tuvo como resultados bordes muy precisos en algunas lesiones, pero no es efectivo en todas las imágenes de ultrasonido, por lo que se han desarrollado alternativas como se muestra a continuación.

3.4. Segmentación por contornos activos

Esta técnica, también conocida como *Snakes*, se ha vuelto cada vez más popular en los últimos años. Estos modelos se basan en propiedades físicas y se utilizan para trazar bordes mediante curvas elásticas que se deforman bajo la acción de fuerzas externas (específicas de la imagen), que son las que hacen que la curva se aproxime a un mínimo local del objetivo de segmentación, o internas (producto del movimiento de la curva misma), que son las que suavizan la curva. Este método mejora la curva inicial minimizando la energía en lugar de detectar automáticamente los bordes de la lesión.

- Energía interna: esta categoría distingue los siguientes elementos:
 - Energía de contorno: Actúa como una fuerza de resistencia y muestra la elasticidad de una curva a que se pliegue o estire.
 - Energía de curvatura: Esta energía representa la rigidez de la curva. Actúa como una fuerza que hace que la curva resista a doblarse y torcerse.
- Energía externa: Se utiliza para acercar la curva al mínimo local del objeto de interés.

Para delinear el borde, simplemente se coloca la curva inicial cerca del borde de interés, que se adaptará gradualmente a la forma de borde deseada (Figura 3.8).

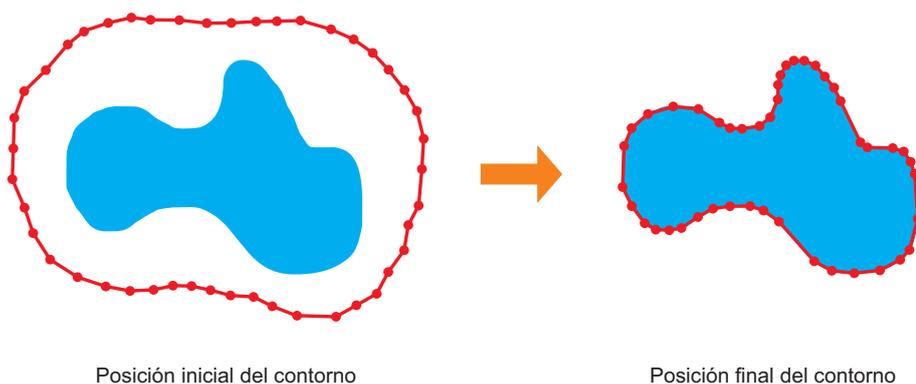


Figura 3.8: Comportamiento de un contorno activo.

Este método de segmentación aplicado a las imágenes de ultrasonido es útil en aplicaciones clínicas porque la segmentación de lesiones es independiente de la experiencia del radiólogo que analiza las imágenes. Basta con establecer la curva inicial para que el médico pueda realizar una segmentación semiautomática, lo que evita errores humanos y hace que el diagnóstico sea más preciso (Figura 3.9), la gran desventaja es que depende de la curva inicial dibujada por el usuario y por tanto no siempre otorga buenos resultados (Figura 3.10).

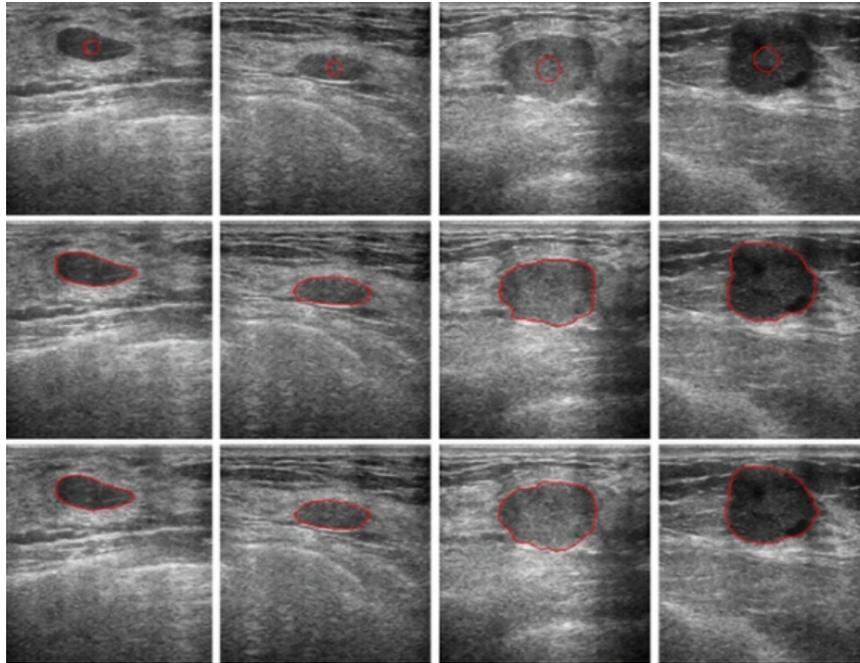


Figura 3.9: Por renglones. Arriba: curva inicial, en medio: seg. experto, abajo: seg. contornos activos [19].

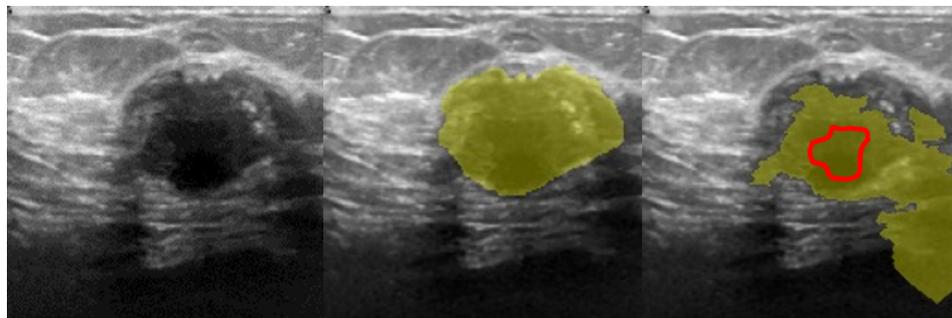


Figura 3.10: Izq: imagen original, centro: seg. experto, der: seg. contornos activos. En rojo se muestra la curva inicial.

3.5. Segmentación con redes neuronales

Los seres humanos tienen la capacidad de reconocer todos los objetos del mundo que los rodea, pero las computadoras por sí solas no pueden realizar esta tarea.

Las *redes neuronales artificiales* (Artificial Neural Networks: ANN) son un grupo de neuronas acomodadas en capas. Estas capas incluyen una capa de entrada que acepta datos, una o más capas ocultas que procesan la entrada y una capa de salida que produce resultados (Figura 3.11).

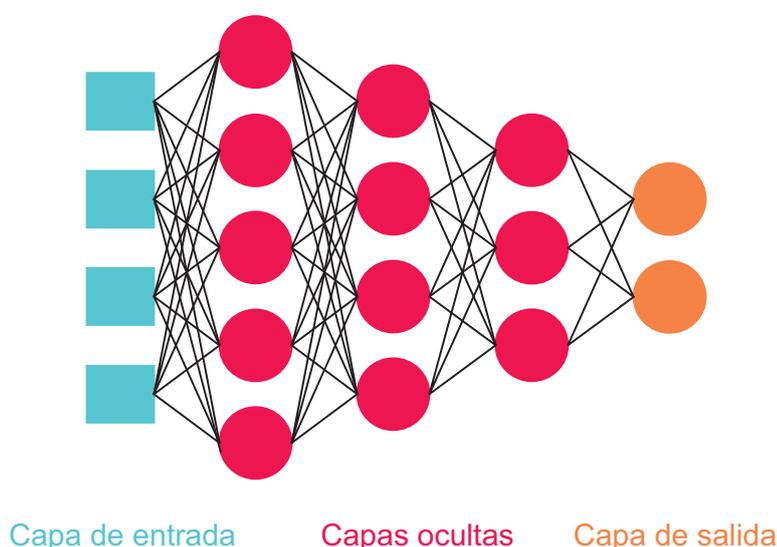


Figura 3.11: Ejemplo de red neuronal artificial.

La salida de una neurona es la activación de la suma ponderada de las entradas (Figura 3.12). Esta activación se refiere a la función de activación ocupada por las ANN, las cuales introducen no linealidades a la red, esto permite que la ANN aprenda funciones complejas entre entradas y salidas. Si una neurona no tiene una función de activación solamente podrá aprender funciones lineales, por lo que no podrá hacer tareas como reconocimiento de objetos ya que éstas son relaciones complejas.

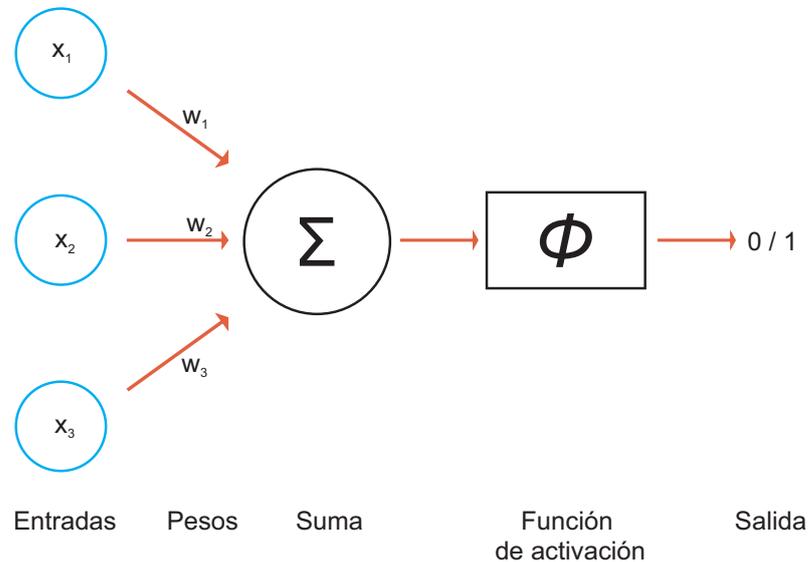


Figura 3.12: Neurona.

Para segmentar una imagen usando una ANN, se necesita convertir la imagen 2D en un vector 1D, lo cual acarrea algunos problemas:

- Se pierde la información espacial de la imagen.
- A medida que aumenta el tamaño de la imagen, el número de parámetros a entrenar aumenta significativamente.

Debido a estos problemas es que las imágenes se procesan utilizando otro método conocido como *redes neuronales convolucionales* (Convolutional Neural Networks: CNN).

Numerosos estudios para comprender el funcionamiento del cerebro humano han demostrado que ciertas partes del cerebro responden a ciertos tipos de estímulos. Por ejemplo, algunas partes de la corteza visual se activan cuando se detectan bordes verticales, otras se activan cuando se detectan bordes horizontales y otras cuando se detectan formas, colores o rostros. Con base en estos estudios, se diseñó un método en el que cada una de las capas funciona como cuando se detecta un estímulo en el cerebro humano. Esta idea dio origen a las redes neuronales convolucionales. Este tipo de redes están constituidas por capas convolucionales.

3. ESTADO DEL ARTE

La capa de convolución es una colección de matrices llamadas filtros que se utilizan para realizar convoluciones en matrices de características como una imagen. Estas convoluciones funcionan de la siguiente manera: un filtro detecta estímulos específicos en la imagen. Este filtro se desplaza por la imagen de entrada y multiplica cada elemento del filtro por la parte de la imagen en la que se encuentra actualmente. Finalmente agrupa los resultados en una sola celda de salida. El filtro repite este proceso para cada parte de la imagen por la que se va desplazando para obtener una matriz de características de la imagen cuando termina (Figura 3.13).

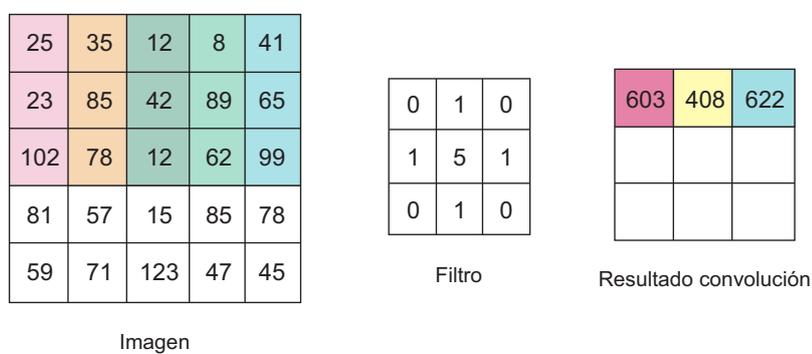


Figura 3.13: Convolución.

Las CNN (Figura 3.14) tienen varias ventajas, que incluyen:

- La red aprende filtros sin tener que ser explícitos en la arquitectura. Estos filtros ayudan a extraer los atributos relevantes de la imagen de entrada.
- Se tiene una conexión local: las neuronas en una capa no están conectadas a todas las neuronas de la capa anterior, sólo a un número pequeño de ellas, por lo que se reducen muchos parámetros.
- Parámetros de entrenamiento compartidos debido a que se aplica el mismo filtro en toda la imagen.
- Este método es muy preciso en la segmentación de lesiones y es ideal para el procesamiento de imágenes.

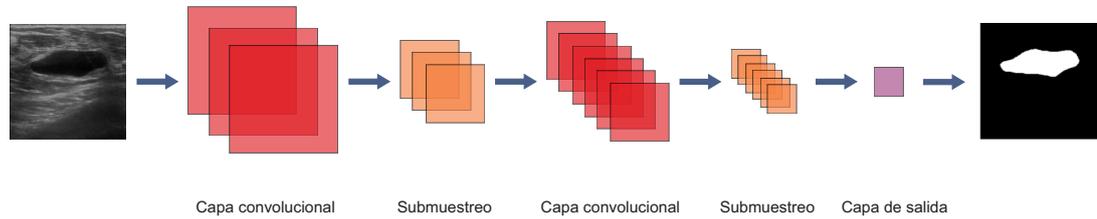


Figura 3.14: Red neuronal convolucional.

Las CNN son capaces de reconocer diferentes objetos en una imagen, pero para ello se entrenan con varios ejemplos para que las capas de la red puedan generalizar la función. Las capas convolucionales se van apilando para poder crear mapas de características en múltiples resoluciones, por lo que los niveles superiores aprenden atributos de campos receptivos cada vez más grandes. Así, la red será capaz tanto de reconocer objetos en imágenes que ya ha visto antes como en ejemplos nuevos.

El proceso de entrenamiento se basa en el proceso de ajuste hacia adelante y hacia atrás, donde se mejora el valor de los pesos y sesgos de conexión entre las capas de neuronas y, conforme se va iterando, estos pesos y sesgos se ajustan al nivel óptimo. Para las CNN, es necesario ajustar los valores de pesos y sesgos para diferentes filtros. La principal ventaja computacional de las CNN es que todos los campos receptivos en una capa tienen pesos y sesgos compartidos, lo que resulta en pocos parámetros entrenables. Para entrenar la red se debe contar con los datos de entrada y sus correspondientes datos de salida deseados.

Un problema común con las CNN es la gran cantidad de datos necesarios para entrenar el modelo. Esto se puede resolver con técnicas de aumento de datos, que permiten generar más datos de entrenamiento a partir de datos existentes.

Algunas de las arquitecturas de CNN más conocidas son: AlexNet [24], VGGNet [35], ResNet [16], GoogLeNet [37] y DenseNet [18].

Metodología

En este capítulo se profundizará en la metodología que se siguió para el desarrollo de esta tesis, en donde se describe una aplicación desarrollada en Swift para la segmentación manual de las imágenes, la adquisición de datos, el pre procesamiento de las imágenes y los métodos propuestos para la reducción de patrón de moteado así como la segmentación automática de lesiones en imágenes de ultrasonido.

4.1. Adquisición de datos

El Laboratorio de Ultrasonido Médico del Instituto de Física, en colaboración con el Área de Radiodiagnóstico del INCan, ha adquirido una colección de imágenes de 1310 imágenes de ultrasonido de mama de 24 pacientes a través del protocolo “Evaluación no invasiva de infiltración inmune en cáncer de mama a través de Radiómica por UC (018/006/RTI) (CEI/1240/17)” en el Instituto Nacional de Cancerología.

Esta colección de imágenes está formada por imágenes de señales de radiofrecuencia sin procesamiento, adquiridos a través de la licencia de investigación Axius Direct del equipo Siemens Acuson S2000. Este equipo estuvo en calidad de préstamo por 6 meses en el INCan.

En el pasado, los investigadores sólo tenían acceso a DICOM o datos de imágenes de escáneres de ultrasonido clínicos. Estos datos tienen una utilidad limitada en aplicaciones de investigación, en gran parte debido al hecho de que son el resultado de una larga serie de algoritmos de procesamiento de señales patentados. Axius Direct URI permite a los investigadores la capacidad de adquirir datos antes en la cadena de procesamiento, por lo que se obtendrán imágenes mucho más ruidosas de lo normal.

Debido a la contingencia sanitaria que empezó a vivir el país desde marzo del 2020, no se pudieron recabar las segmentaciones por parte de las radiólogas de las 1310 imágenes. Se empezó a acudir al INCan el 5 de marzo del 2020, viéndose interrumpidas las visitas al hospital ese mismo mes, por lo que solamente se lograron recabar 71 trazos de dos radiólogas. Para solucionar este problema al aplicar la red neuronal, se hizo uso de una biblioteca de Python para realizar aumento de datos, tanto la red como la biblioteca se explicarán más adelante.

4.2. Preparación de datos

El conjunto de imágenes originales venía separado en carpetas por pacientes y cada imagen se identificaba por: *No. lesión, corte, fecha, hora*. Para conservar la confidencialidad de los pacientes, se desarrolló un programa que acoplara todas las imágenes en la misma carpeta de forma intercalada para que las lesiones de cada paciente quedaran separadas. Posteriormente, el programa va renombrando cada imagen con números consecutivos y cambiando su fecha de modificación, esto para poder tener un orden al cargar las imágenes en la aplicación para iPad que se explicará más adelante. Finalmente, el programa registra en un archivo .xlsx el nombre original de la imagen así como su nuevo identificador para no perder las referencias iniciales.

Posteriormente se creó otro programa que ayuda a quitar el borde de las imágenes originales automáticamente (Figura 4.1). Una vez que ya se tienen las imágenes procesadas, se puede proceder a la recabación de segmentaciones manuales por parte de los radiólogos.

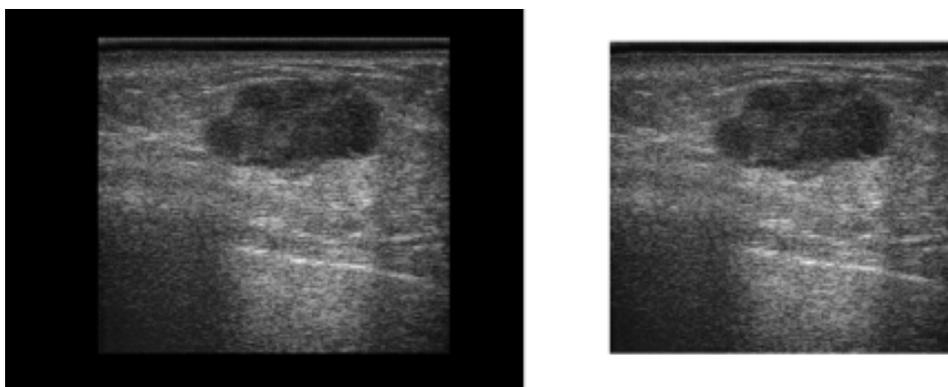


Figura 4.1: Izq: imagen con borde, der: imagen procesada

Para poder acudir al INCan a realizar las segmentaciones manuales, se seleccionaron imágenes de 10 pacientes y se cargaron en el iPad. Como se explicó anteriormente, todas las imágenes tienen una fecha de modificación distinta, de esta manera se pueden cargar de forma ordenada en el dispositivo. Lo anterior es muy importante para poder identificar las imágenes una vez que estén cargadas en el iPad, ya que estos dispositivos no mantienen el nombre original de las imágenes y si no se tuviera un orden específico se perderían las referencias totalmente. Esto ayudará a identificar todas las imágenes una vez que se tengan las segmentaciones manuales y se tengan que procesar en la computadora.

4.3. Aplicación para la segmentación manual

En el 2019, durante la realización de mi servicio social de licenciatura en el programa de trabajo “Investigación Instituto de Física” número 2019-12/82-921 en el Instituto de Física de la UNAM, desarrollé una aplicación para iPad en el lenguaje de programación Swift para hacer del marcado de las lesiones en imágenes de ultrasonido una tarea más sencilla para las radiólogas (A.1).

La aplicación consiste en obtener todas las imágenes de la galería del iPad para cargarlas en la galería de la aplicación, cuando es la primera vez que se ejecuta la aplicación en el dispositivo, se pide permiso para acceder a las imágenes (Figura A.1). Posteriormente, si se acepta el permiso, se puede acceder a cada imagen para editarla y que los radiólogos puedan hacer la segmentación manual.

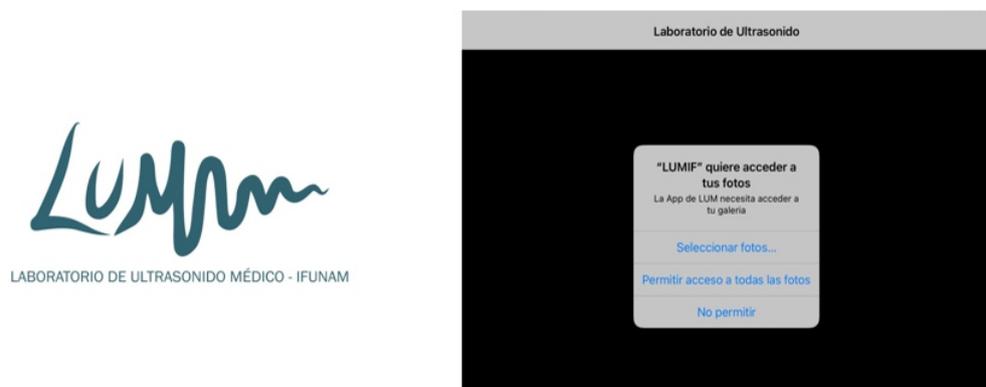


Figura 4.2: Izq: inicio de aplicación, der: permiso para acceder a galería

4. METODOLOGÍA

En la vista de edición, en la parte de abajo de la pantalla se muestra si esa imagen ya fue marcada o aún no, si no, se muestra el mensaje “Segmentación no realizada” en rojo, y si sí, se muestra el mensaje “Segmentación guardada” en verde (Figura 4.3).

Por otro lado, se tienen los botones de “Cancelar”, “Limpiar” y “Guardar”; con el primero se cancela la edición, con el segundo se limpian todos los trazos que se hayan realizado sobre la imagen y con el tercero se guarda la segmentación y se pasa automáticamente a la siguiente imagen (Figura 4.3). Una vez que se sale de la edición, con el botón “Regresar” se puede acceder nuevamente a la galería (Figura 4.4).

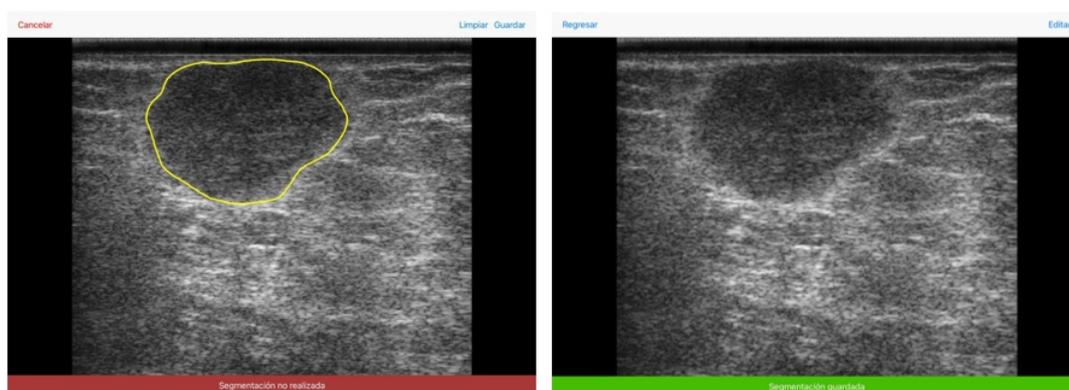


Figura 4.3: Izq: imagen en edición, der: imagen guardada

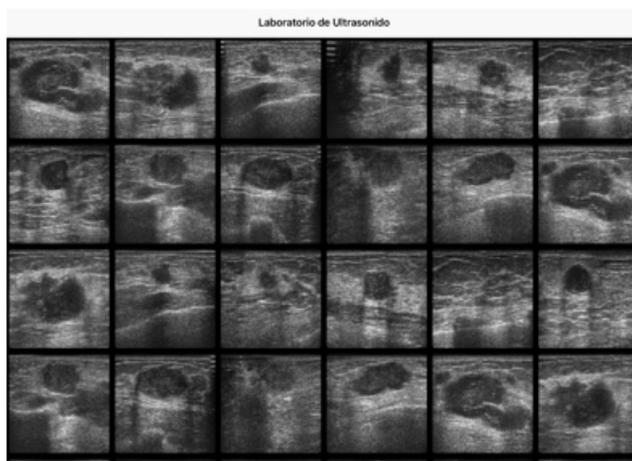


Figura 4.4: Galería de la aplicación

Para esta tesis se realizaron modificaciones a la aplicación, las cuales consisten en:

- Agregar un botón “Sin lesión” para aquellas imágenes en las que las radiólogas no encuentren lesiones (Figura 4.5).
- Eliminar la confirmación del botón “Guardar” para que el marcado pueda ser más rápido.
- Cerrar automáticamente el trazo si el radiólogo no llega a cerrarlo para poder procesar la máscara más rápidamente.
- Marcar en la galería aquellas imágenes que ya fueron segmentadas (Figura 4.6).
- Mandar una alerta cuando se llega al final de las imágenes que indica que no hay más imágenes por marcar (Figura 4.6).



Figura 4.5: Izq: botón “Sin lesión” presionado, der: muestra de lo que se guarda al presionar ese botón.

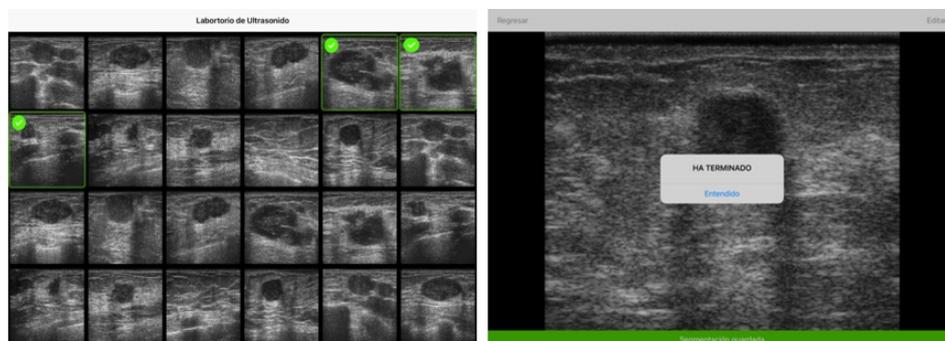


Figura 4.6: Izq: galería que muestra las imágenes que ya se marcaron, Der: alerta de terminación.

4. METODOLOGÍA

Se obtienen dos resultados por imagen de esta aplicación, el primero es el trazo en un fondo transparente, el cual posteriormente se procesará para convertirlo en la máscara y el segundo es el trazo sobre la imagen de ultrasonido. Ambos resultados poseen un índice en la parte superior izquierda, el cual ayuda a identificar las imágenes una vez que se traspasen a la computadora, ya que los índices corresponden al nombre de las imágenes cuando se prepararon los datos (Figura 4.7). Cuando se presiona el botón “Sin lesión” como se mostró anteriormente, en las dos imágenes resultantes se presenta la leyenda “No se encontró lesión en esta imagen”, esto para tener la referencia una vez que se procesen en la computadora (Figura 4.8).

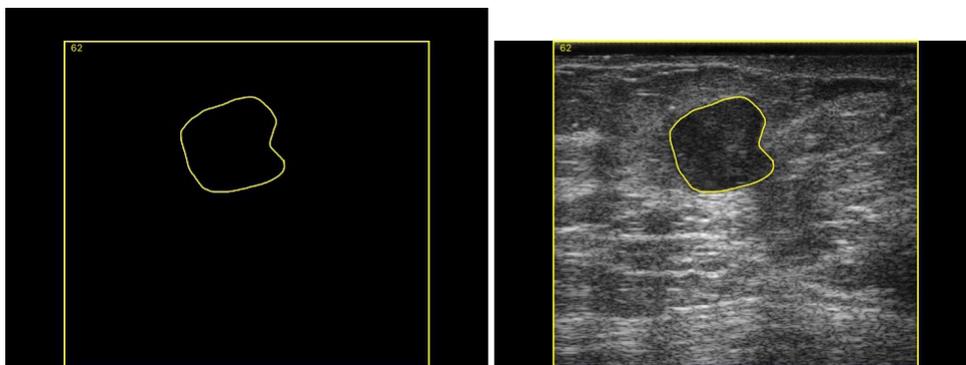


Figura 4.7: Resultados de la segmentación manual.

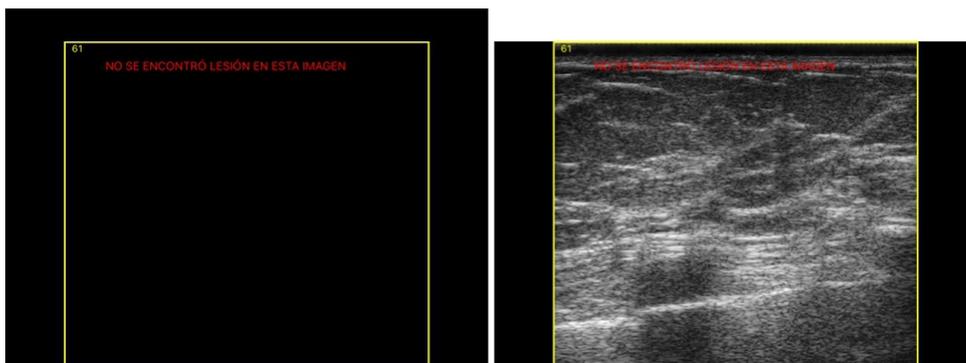


Figura 4.8: Resultados cuando se presiona botón “Sin lesión”

4.4. Métodos propuestos

En las siguientes secciones se presentan los métodos que se proponen para llevar a cabo la reducción de patrón de moteado y la segmentación de lesiones en imágenes de ultrasonido.

4.4.1. Reducción de patrón de moteado en imágenes de investigación

El querer reducir el patrón de moteado de las imágenes de investigación es debido a que en la primera visita al INCa, las radiólogas que segmentaron las imágenes manualmente señalaron que estaban sumamente ruidosas para ellas ya que suelen trabajar con imágenes clínicas (limpias) (Figura 4.9). Como se ha explicado antes, una buena segmentación depende de la experiencia que tenga el radiólogo que analiza la imagen, si los radiólogos no están acostumbrados a cierto tipo de imágenes se podrían cometer errores en el marcado de las lesiones.

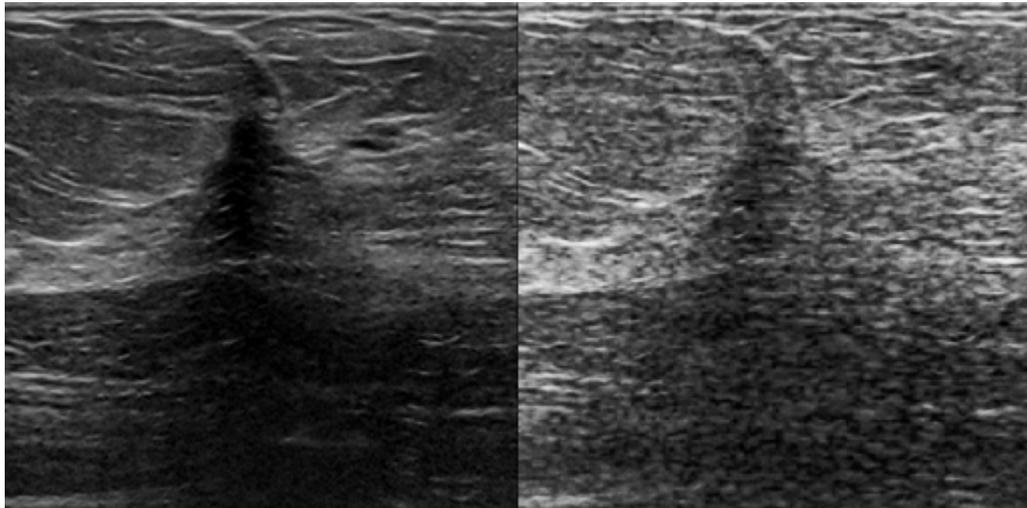


Figura 4.9: Izq: Imagen clínica, der: imagen de investigación.

La técnica que se propone en este trabajo para reducción del ruido es el uso de autocodificadores.

4.4.1.1. Autocodificadores

Los autocodificadores son un tipo de red neuronal que pueden aceptar una colección de imágenes como entrada, comprimir los datos para tener una representación más

4. METODOLOGÍA

pequeña de ellos y reconstruir la entrada desde los datos comprimidos. Los autocodificadores tienen dos componentes, un codificador que recibe la entrada y la comprime, y el decodificador que toma la versión comprimida de los datos y la reconstruye.

Las aplicaciones de los autocodificadores van desde la reducción de dimensionalidad, reducción de ruido, detección de anomalías, super resolución de imágenes, etc.

En el caso de reducción de ruido, lo que hacen los autocodificadores es entrenar el filtro de eliminación de ruido basado en el conjunto de datos con el que se está trabajando. Al controlar la entrada y los valores objetivo antes de comenzar el proceso de entrenamiento, es posible que se pueda entrenar el codificador y decodificador de manera que elimine el ruido. Para esto se deben proporcionar imágenes ruidosas a la entrada e imágenes limpias a la salida; así, codificador y decodificador aprenderán a eliminar el ruido de las imágenes ya que se debe producir como salida la imagen sin ruido (Figura 4.10).

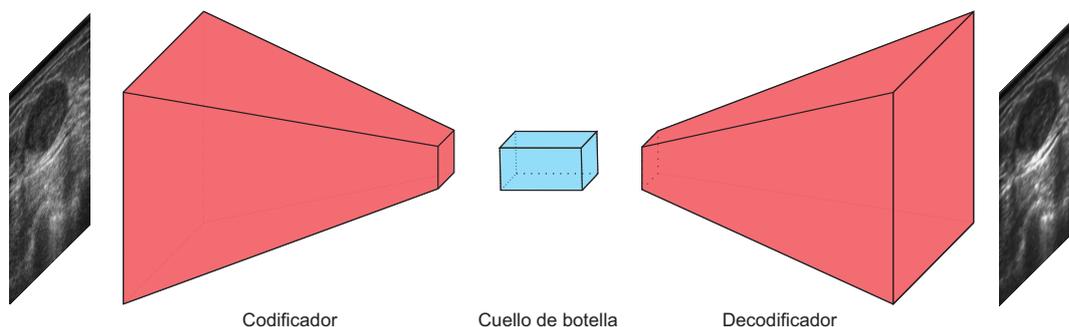


Figura 4.10: Entrenamiento de autocodificador para reducción de ruido.

Para poder entrenar al autocodificador, se utilizará una colección de imágenes tomada de [5] en la que se encuentran 780 imágenes clínicas de ultrasonido de mama sin lesiones, con lesiones benignas y con lesiones malignas. Estas imágenes están filtradas y no tienen tanto ruido como las imágenes de investigación con las que se cuenta para este trabajo.

Para poder simular las imágenes de investigación a partir de las imágenes clínicas, primero se creó una imagen (Figura 4.11) que representa el fondo ruidoso que se genera en las imágenes de investigación con ayuda del programa *Field II* [21], que es un software que se ejecuta en MATLAB capaz de simular imágenes de ultrasonido.

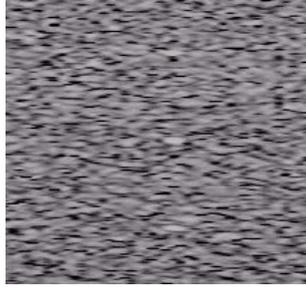


Figura 4.11: Imagen de ruido *speckle* creada con *Field II* [21].

Posteriormente se obtuvo la varianza (0.0316) de la imagen para poder agregarles ruido a las imágenes clínicas por medio de la función *random_noise* de la biblioteca *skimage* de Python. De esta forma se tiene el par de *datasets* que se necesitan para entrenar el autocodificador: *dataset* con imágenes de investigación simuladas y otro *dataset* con su respectiva imagen clínica (Figura 4.12).

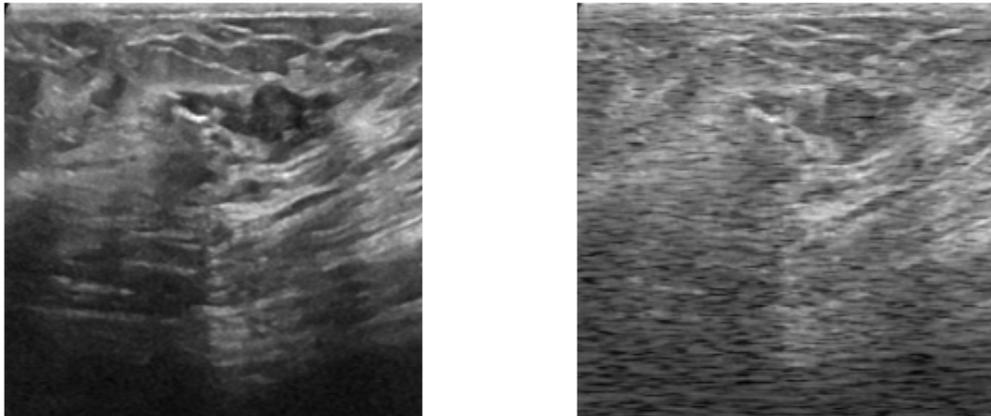


Figura 4.12: Izq: imagen clínica, der: imagen de investigación simulada.

Para poder asegurar que las imágenes creadas efectivamente tienen la calidad de las imágenes de investigación, se utilizó la relación señal ruido (SNR), la cual describe la nitidez que tiene una imagen. Si una imagen es muy nítida la SNR será grande, mientras que si la nitidez es baja, la SNR será pequeña.

Para realizar esta medición se utilizaron dos ventanas de 20×20 , una para seleccionar la parte central de la lesión (A) y otra para seleccionar una parte homogénea del fondo (B) (Figura 4.13). Después, siguiendo la ecuación 4.1 se obtuvo la SNR de cada imagen [28].

4. METODOLOGÍA

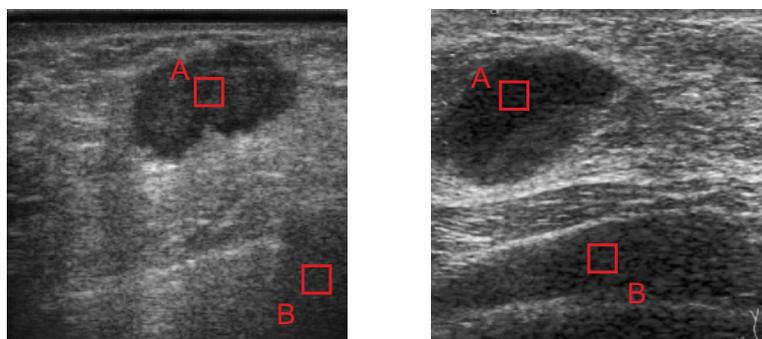


Figura 4.13: Izq: imagen de investigación, der: imagen de investigación simulada.

$$SNR = 20\log_{10} \left(\frac{mean(I_A)}{std(I_B)} \right) \quad (4.1)$$

Para obtener una SNR promedio se utilizaron 50 imágenes de investigación y 50 imágenes de investigación simuladas. Como se puede observar en la siguiente tabla, la SNR de las imágenes de investigación simuladas es más pequeña, lo que nos indica que las imágenes creadas tienen menor calidad que las imágenes de investigación originales.

	Imágenes de investigación	Imágenes de investigación simuladas
SNR	7.3453 dB	6.9041 dB

Para este trabajo se construyó un autocodificador que tiene 4 bloques en el codificador y 4 bloques en el decodificador. Los bloques del codificador se componen por dos capas convolucionales con función de activación ReLU y una capa de *max pooling*. Las capas convolucionales del primer bloque tienen 16 filtros, del segundo bloque 32 filtros, del tercer bloque 64 filtros y del cuarto bloque 128 filtros. El cuello de botella cuenta con dos capas convolucionales con 256 filtros y función de activación ReLU.

Por otro lado, los bloques del decodificador se componen por una capa de *upsampling* y dos capas convolucionales con función de activación ReLU. Las capas convolucionales del primer bloque tienen 128 filtros, del segundo bloque 64 filtros, del tercer bloque 32 filtros y del cuarto bloque 16 filtros.

La capa de salida está formada por una capa de convolución y función de activación

lineal. Todas las capas convolucionales utilizan filtros de 3×3 .

Las capas utilizadas se explicarán a continuación.

- **Capas convolucionales:** Como se explicó anteriormente, las capas convolucionales aplican filtros a las imágenes de entrada para crear mapas de características que resuman las características de entrada. Apilar capas convolucionales en redes profundas permite que las capas cercanas a la entrada aprendan características de bajo nivel como líneas verticales y horizontales, y las capas más profundas aprenden características más abstractas, como formas u objetos específicos.
- **Capa *max pooling*:** Una capa de *pooling* o agrupación se agrega después de haber aplicado una no linealidad a una capa convolucional. Esta capa ayuda a tener un mapa de características agrupado, de tal forma que si se presentan pequeñas variaciones en la ubicación de la característica de la imagen de entrada, la agrupación hará que la característica se encuentre en la misma ubicación.

Para realizar una agrupación se tiene que elegir el tamaño del filtro, el cual en redes neuronales convolucionales regularmente se ocupa de 2×2 con un tamaño de paso de 2 píxeles. Lo anterior hace que se produzca un mapa de características que será de un cuarto del tamaño de la entrada, ya que cada dimensión se reducirá a la mitad.

Existen diferentes operaciones para agrupar las características, como la agrupación promedio o *average pooling*, en la que se calcula un valor promedio para cada filtro en el mapa de características. La que se ocupa en este trabajo es la agrupación máxima o *max pooling*, en la que se selecciona el valor más alto de cada filtro en el mapa de características (Figura 4.14).

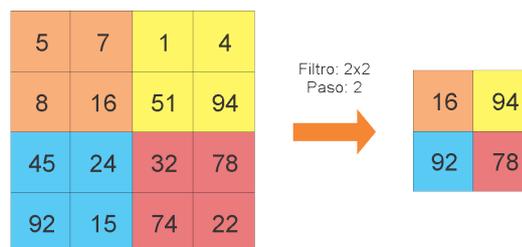


Figura 4.14: Capa *max pooling*.

4. METODOLOGÍA

- Capa de *upsampling*: Contrario a la capa de *pooling*, la capa de *upsampling* es una técnica que expande la dimensión de la imagen a un mapa de características del tamaño deseado. En este caso, la capa de *upsampling* utilizará la técnica de vecino más cercano, por lo que copiará los elementos de la matriz tantas veces como se indique. Por ejemplo, si se utiliza una ventana de 2×2 , cada pixel será copiado 4 veces, haciendo que el tamaño de la imagen crezca al doble (Figura 4.15).

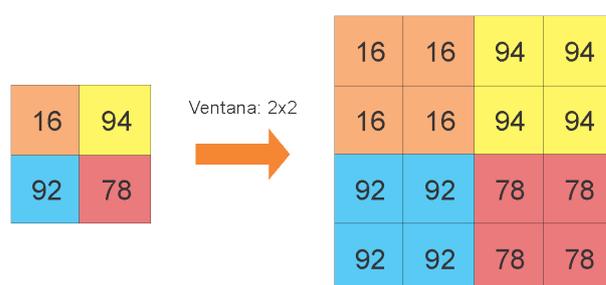


Figura 4.15: Capa *upsampling*.

Las entradas del autocodificador implementado son de tamaño 256×256 , las cuales se reducen hasta 16×16 (Figura 4.16). El optimizador utilizado es Adam y la función de pérdida utilizada es el error cuadrático medio, la cual mide qué tan parecida es la reconstrucción de la entrada contra la imagen original.

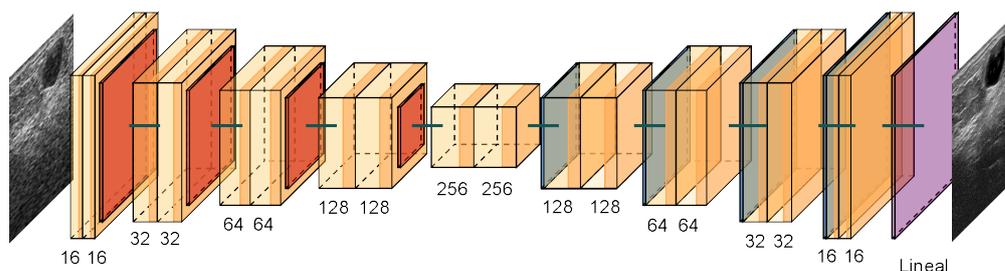


Figura 4.16: Modelo utilizado.

4.4.2. Segmentación de lesiones en imágenes de investigación

Como se mencionó anteriormente, en los últimos 20 años se han propuesto muchas alternativas de segmentación automática y semiautomática para lesiones mamarias, pero se ha detectado que no pueden segmentar masas o tumores muy pequeños, y detectar tumores pequeños es importante para encontrar cánceres en etapa temprana.

Debido a que varias arquitecturas de redes neuronales convolucionales han tenido buenos resultados en segmentación de imágenes de ultrasonido, es que se desarrolla la arquitectura STAN [33], la cual muestra que puede segmentar lesiones de diferente tamaño.

4.4.2.1. Redes neuronales convolucionales: Small Tumor Aware Network (STAN)

La arquitectura STAN está basada en la arquitectura U-Net [20], la cual fue desarrollada para la segmentación de imágenes biomédicas. La arquitectura U-Net (Figura 4.17) tiene dos secciones. La primera es la parte del codificador, que se utiliza para capturar las características principales de la imagen. El codificador está compuesto por una pila de capas convolucionales y de *max pooling*. La segunda sección es la parte del decodificador, que se utiliza para permitir una localización precisa de los objetos mediante convoluciones transpuestas.

Cada bloque en el codificador tiene dos capas de convolución con filtros de 3×3 seguidas de una capa *max pooling* de 2×2 . En cada nivel de la arquitectura se va duplicando el número de filtros utilizados en las capas de convolución, esto para que la arquitectura pueda aprender las estructuras complejas de forma efectiva.

En el cuello de botella, que es la capa en la que se unen codificador y decodificador, se tienen dos capas de convolución con filtro de 3×3 seguida de una capa de deconvolución con un paso de 2.

Por otro lado, cada bloque en el decodificador tiene dos capas de convolución y una capa de deconvolución. A cada entrada de los bloques se le concatenan los mapas de características del nivel correspondiente en el codificador. Estas son las llamadas conexiones de salto, ya que conectan el codificador y decodificador en su respectivo nivel, lo que asegura que las características aprendidas en la codificación se usen al reconstruir la imagen.

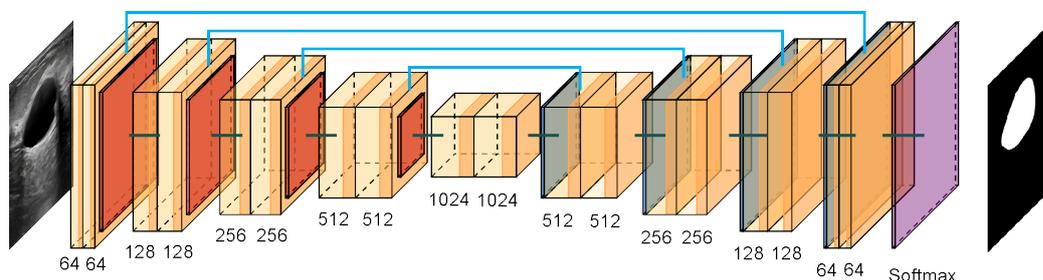


Figura 4.17: Arquitectura U-Net.

Por otro lado, la arquitectura STAN se basa en dos observaciones clave: el tamaño de los tumores varía drásticamente entre pacientes y las redes profundas existentes usan tamaños fijos de los filtros y no pueden detectar tumores de distintos tamaños, por lo que STAN implementa la construcción de mapas de características usando filtros de 3 tamaños diferentes en cada capa del codificador, con la finalidad de que se capturen características en escalas diferentes. Así, en lugar de tener un solo mapa de características, se tienen 3, por lo que se tiene más información.

Al igual que la arquitectura U-Net, la arquitectura STAN (Figura 4.18) contiene una etapa de codificación y otra de decodificación, así como conexiones de salto para recuperar detalles de la imagen en el decodificador, la diferencia es que en el codificador se tienen dos ramas en lugar de una, y se comparten 2 conexiones de salto con el decodificador.

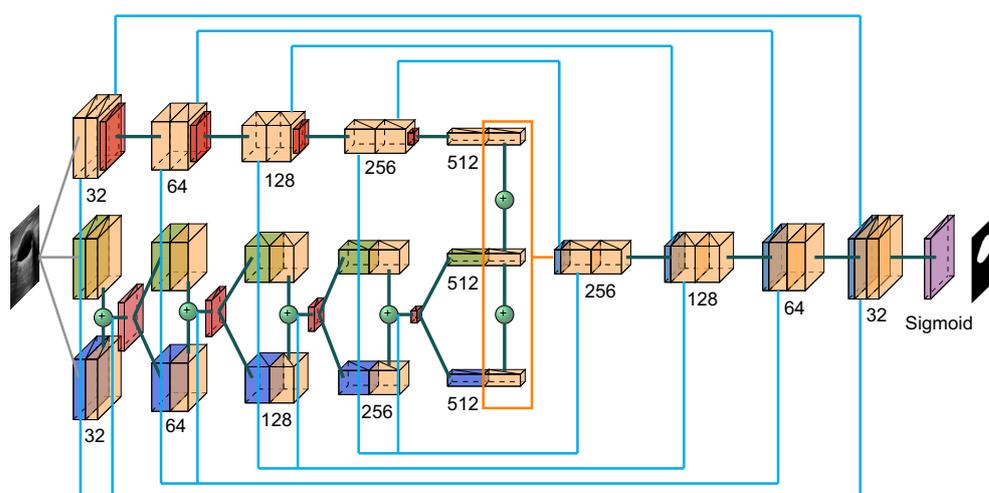


Figura 4.18: Arquitectura STAN.

Cada bloque del codificador se define por tener en la primera rama dos capas convolucionales con un tamaño de filtro igual a 3 y una capa de *max pooling*. Para la segunda rama se tienen dos capas convolucionales, la primera con filtros de 5×5 y de 1×1 seguidos de una capa convolucional con tamaño de filtro igual a 3. Esta última capa se concatena y se pasa por una capa de *max pooling*. Por otro lado, se tienen 2 conexiones de salto que irán al decodificador, la primera conexión sale de la segunda capa de la primera rama y la segunda es una concatenación que sale de la primera capa de la primera rama con la concatenación de capas de la segunda rama.

Para el bloque central ya sólo existe una rama y se define por dos capas convolucionales, la primera aplica filtros de 3×3 , 5×5 y 1×1 , todos estos son seguidos de la segunda capa convolucional que solamente aplica filtros de 3×3 y todas estas salidas se concatenan.

Los bloques del decodificador están definidos por una capa que deconvoluciona la entrada concatenada con la primera conexión de salto del codificador en su respectivo nivel con un filtro de 3×3 . Le sigue una capa que convoluciona filtros de 3×3 con la entrada de la capa concatenada a la segunda conexión de salto en su respectivo nivel del codificador y se aplica otra capa convolucional con un filtro de 3×3 . Finalmente se tiene la capa de salida que es una capa convolucional con un filtro de 3×3 y una función de activación sigmoide.

Esta red ha superado varias arquitecturas de redes neuronales convolucionales, ya que puede detectar tumores de distinto tamaño con una tasa baja de falsos negativos, lo cual es importante para ayudar a detectar enfermedades en sus primeras etapas.

4.4.2.2. Aumento de datos

Debido a que se cuenta con muy pocas imágenes de investigación, se realizará aumento de datos, que sirve para crear copias de las imágenes originales con alteraciones o cambios, de tal forma que ayude a mejorar la capacidad de las redes para detectar objetos.

Para lo anterior se utilizará una biblioteca de python llamada *imgaug*, la cual brinda varias técnicas de aumento como agregar distintos tipos de ruido, desenfoque, brillo, hacer escalamiento, rotaciones, traslaciones, recortes, etc. Estas transformaciones pue-

4. METODOLOGÍA

den combinarse y ejecutarse en orden aleatorio, lo que hace a esta biblioteca muy útil y proporciona una forma muy sencilla de realizar el aumento de datos.

La siguiente imagen muestra un ejemplo de la forma en que funciona *imgaug*. Para esta imagen se aplican recortes, se voltean las imágenes horizontalmente, se agrega ruido y desenfoque, y además cambia el contraste y el brillo. Como se mencionó anteriormente, todas las transformaciones se realizarán de manera aleatoria, lo que resultará en una gran variedad de imágenes a partir de una sola.

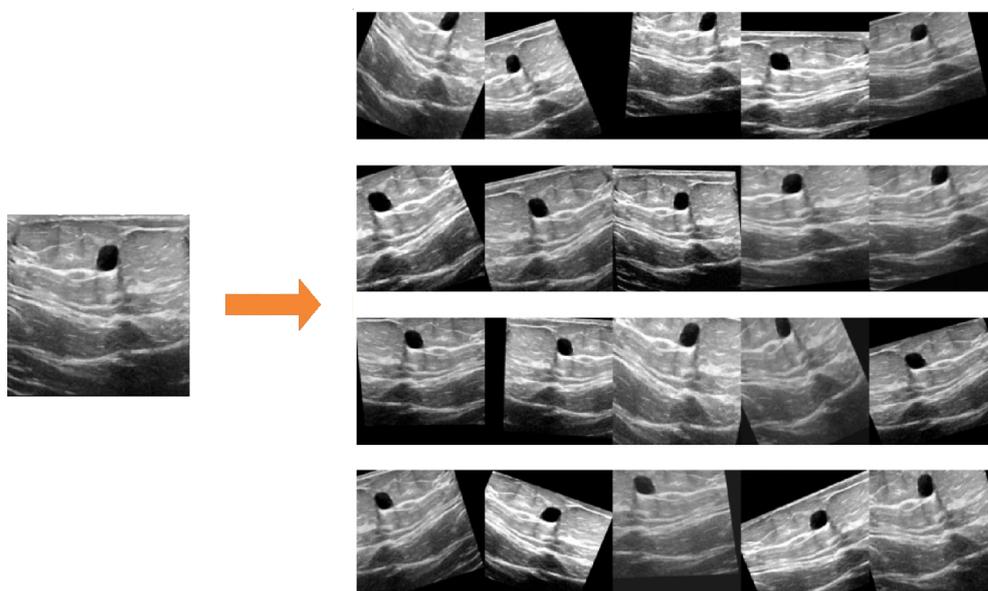


Figura 4.19: Aumento de datos

4.4.3. Validación de los métodos

Para validar el primer método propuesto que es el desarrollo del autocodificador, se utilizará el índice de similitud estructural (SSIM), que es una métrica que cuantifica la degradación de la calidad de la imagen causada por el procesamiento. La idea básica de SSIM es evaluar la similitud de dos imágenes (x y y) a través de los siguientes aspectos:

- Luminancia: Fenómeno por el cual las distorsiones de la imagen tienden a ser menos visibles en regiones brillantes.

$$l(x, y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1}$$

donde:

μ_x es la media de x

μ_y es la media de y

$c_1 = (k_1 L)^2$ para evitar inestabilidad cuando el denominador se acerca a cero.

L es el rango dinámico de píxeles, por ejemplo $[0,255]$

$k_1 = 0.01$ por defecto.

- **Contraste:** Fenómeno por el cual las distorsiones se vuelven menos visibles donde hay textura en la imagen.

$$c(x, y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2}$$

donde:

σ_x^2 es la varianza de x

σ_y^2 es la varianza de y

$c_2 = (k_2 L)^2$ para evitar inestabilidad cuando el denominador se acerca a cero.

$k_2 = 0.03$ por defecto.

L es el rango dinámico de píxeles, por ejemplo $[0,255]$

- **Estructura:** La información estructural es la idea de que los píxeles tienen fuertes interdependencias, especialmente cuando están espacialmente cerca. Estas dependencias llevan información importante sobre la estructura de los objetos en la escena visual.

$$s(x, y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3}$$

donde:

σ_{xy} es la covarianza de x y y

$c_3 = c_2/2$

SSIM es entonces una combinación ponderada de esas medidas comparativas:

$$SSIM(x, y) = [l(x, y)^\alpha \cdot c(x, y)^\beta \cdot s(x, y)^\gamma]$$

α, β, γ representan la importancia de cada métrica. Si todas las métricas se establecen en 1, la fórmula se puede reducir a la siguiente forma:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y+c_1)(2\sigma_{xy}+c_2)}{(\mu_x^2+\mu_y^2+c_1)(\sigma_x^2+\sigma_y^2+c_2)}$$

En este caso se le dará distinta importancia a cada métrica, ya que lo que se quiere evaluar es cuánta estructura se conserva al procesar las imágenes con el autocodificador, por lo que $\gamma = 2$. Se le otorga mayor importancia a la estructura debido a que es necesario conservarla para que los radiólogos puedan hacer una detección y segmentación exitosas.

Por otro lado, también se evaluará la SNR en las imágenes predichas para saber qué tan nítidas son con respecto a las imágenes de investigación creadas y poder comparar la SNR contra las imágenes clínicas originales.

Las imágenes que se compararán son las salidas predichas por la red contra las imágenes clínicas objetivo. Idealmente, estas imágenes deberían ser iguales ($SSIM = 1$), por lo que se espera que el SSIM sea lo más cercano a 1 posible, así como se espera que la SNR sea parecida a la SNR de las imágenes clínicas.

Por otro lado, para evaluar la arquitectura STAN se llevarán a cabo varios pasos:

1. Se replicará la arquitectura STAN y se compararán los resultados con lo que muestran en [33] para asegurar que la réplica construida es tan buena como la presentada en el artículo.
2. Se medirá el desempeño del primer modelo utilizando las imágenes procesadas con el autocodificador y realizando aumento de datos.
3. Se medirá el desempeño del segundo modelo que se inicializará con los pesos y sesgos del primer modelo. Se entrenará con el *dataset* de las imágenes procesadas con el autocodificador y el *dataset* con las imágenes de investigación originales utilizando aumento de datos.

4. Para poder medir si en realidad los resultados que entrega la red son aceptables o no, se calculará el error entre la segmentación de dos radiólogas del INCan y se comparará con los resultados de la red.

Las métricas de evaluación que se utilizarán son:

- Sensibilidad: probabilidad de que un sujeto enfermo efectivamente esté enfermo.
- Precisión: indica qué proporción de predicciones como enfermo fueron correctas.
- Especificidad: probabilidad de que los casos sanos estén correctamente identificados.
- Exactitud: probabilidad de clasificar correctamente los elementos.
- El índice Dice (DSC): calcula qué tanto se parece el área segmentada por la red con el área segmentada por el experto, dándole más peso a los elementos comunes entre ambas segmentaciones.
- La tasa de error del área (AER): indica qué tan probable es que el área predicha no sea el área segmentada por el experto.

En este capítulo se presentan los resultados obtenidos de la reducción de patrón de moteado, tres experimentos diferentes con la arquitectura utilizada para la segmentación automática de lesiones, las métricas utilizadas para calcular el error de segmentación entre radiólogas, la comparación del desempeño de los modelos utilizados con el error de segmentación entre radiólogas y finalmente la discusión de resultados.

5.1. Máscara de segmentación

A partir de las segmentaciones manuales realizadas por las radiólogas del INCAN, se desarrolló un programa en MATLAB, el cual se puede encontrar en [7], que convirtiera los trazos con transparencias en las máscaras que se utilizarán para todos los algoritmos.

Este programa consiste en encontrar los bordes reales de la imagen marcados con amarillo (Figura 5.1) y copiarlos a una nueva matriz de datos. Una vez que se copia la imagen real, se rellena la lesión que esté marcada, para que posteriormente se elimine el número que se tiene como referencia en la parte superior izquierda así como el borde, para que de esta manera esté terminada la máscara (Figura 5.2).

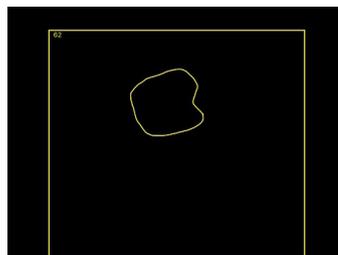


Figura 5.1: Ejemplo de imagen de salida de la aplicación.

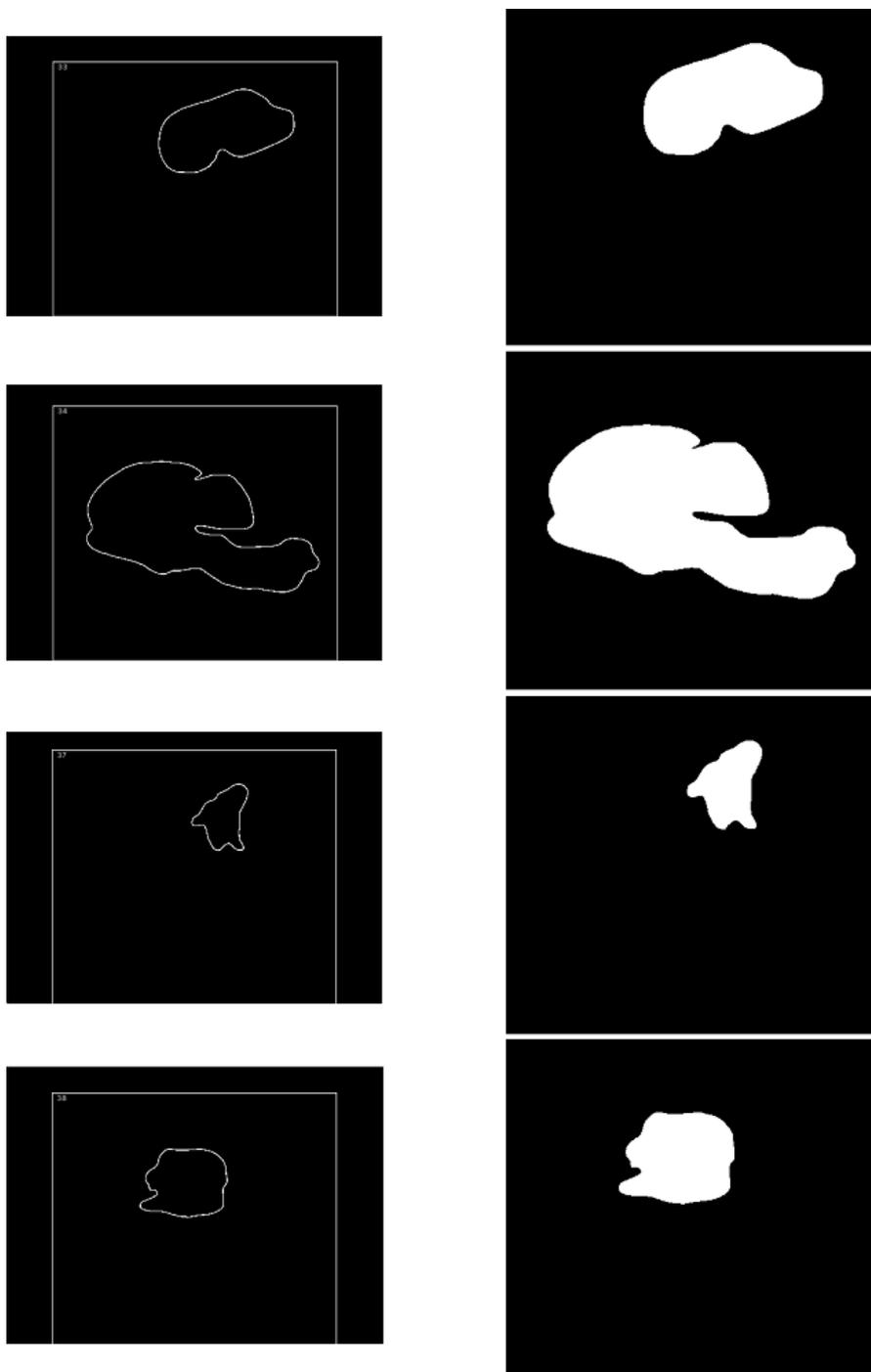


Figura 5.2: Izq: Trazos que guarda la aplicación, der: máscaras construidas.

5.2. Reducción de patrón de moteado en imágenes de investigación

El modelo del autocodificador explicado anteriormente se ejecutó por 100 épocas, teniendo una pérdida baja desde la época 40, con un *batch size* = 16.

El comportamiento de la pérdida durante el entrenamiento del autocodificador se muestra en la siguiente gráfica:

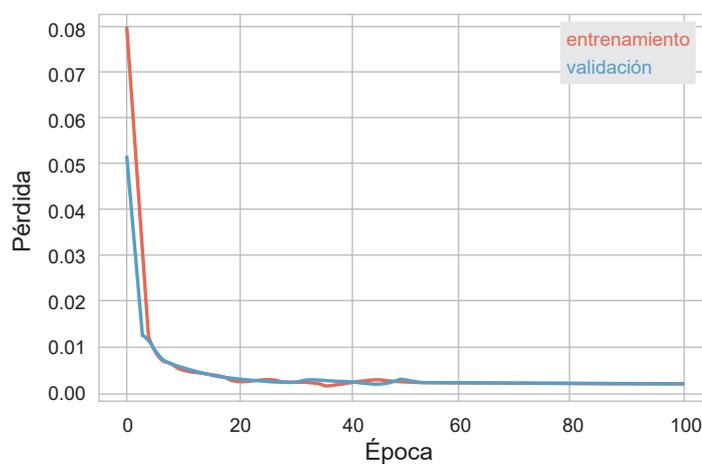


Figura 5.3: Comportamiento de la pérdida en autocodificador.

Como se explicó anteriormente, el desempeño de este modelo se mide con el índice de similitud estructural (SSIM), y mientras más cercano a 1 sea, significa que la imagen predicha es muy parecida a la imagen objetivo. Además, se midió el SSIM de las imágenes de investigación simuladas contra las imágenes clínicas, así como la SSIM de las imágenes de investigación simuladas contra las imágenes predichas.

	Investigación simuladas vs clínicas	Investigación simuladas vs predichas	Imágenes objetivo vs predichas
SSIM	0.7506	0.7473	0.9147

Estos resultados indican que las imágenes clínicas son 75 % parecidas a las imágenes de investigación simuladas debido a que se pierde estructura por el ruido que se agregó. De la misma manera sucede con las imágenes de investigación simuladas contra las pre-

5. RESULTADOS

dichas, mientras que las imágenes predichas son 91 % parecidas a las imágenes objetivo, lo que indica que efectivamente se está realizando la reducción de patrón de moteado y conservándose la estructura de la mama. Se muestran a continuación 5 ejemplos del conjunto de prueba:

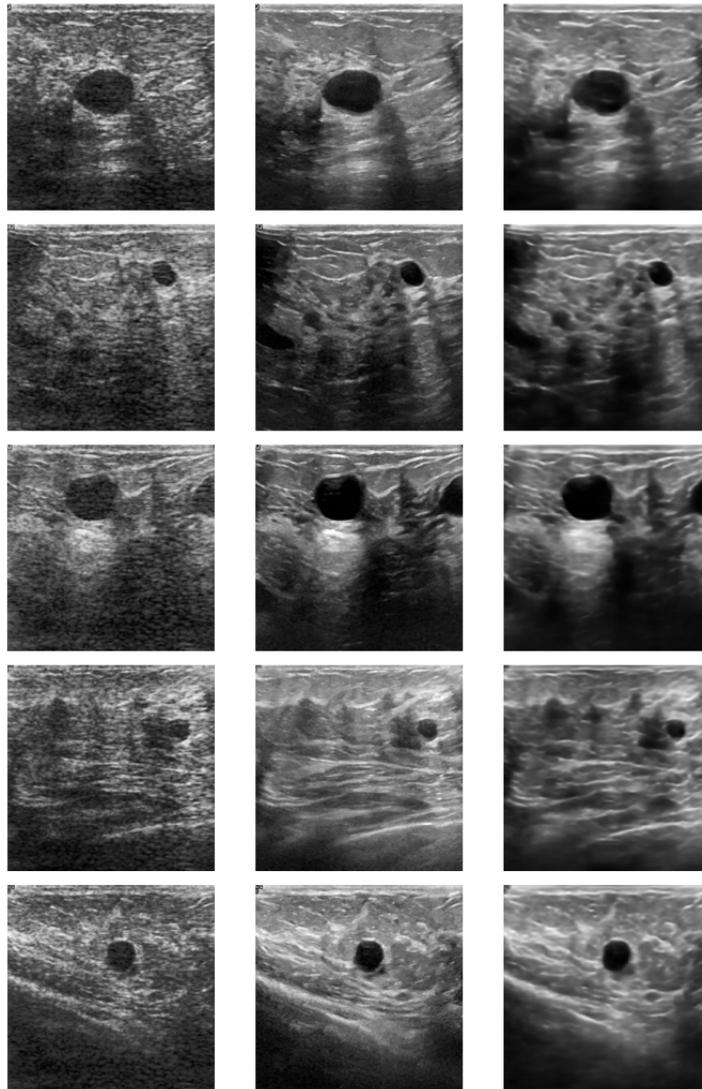


Figura 5.4: Izq: imagen ruidosa, centro: imagen objetivo, der: imagen predicha con autocodificador.

5.2 Reducción de patrón de moteado en imágenes de investigación

Una vez que se obtuvieron los resultados deseados en el conjunto de validación de las imágenes clínicas, se probó el método con las imágenes de investigación:

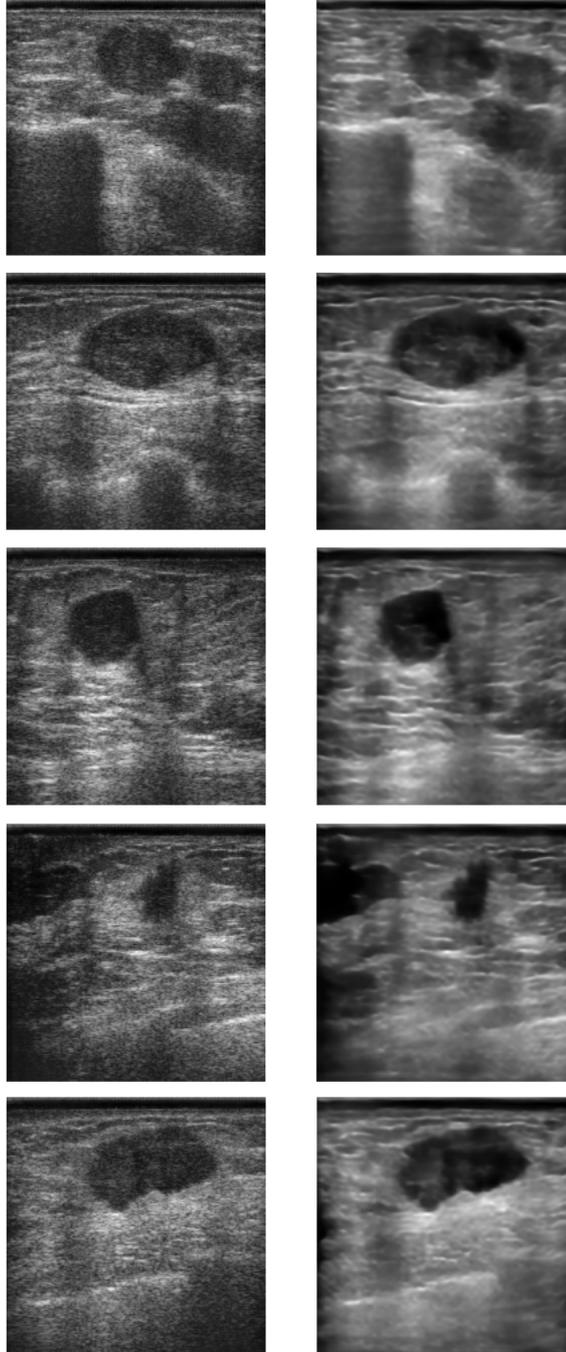


Figura 5.5: Izq: imagen de investigación, der: imagen predicha con autocodificador.

Además, se midió la SNR de las imágenes resultantes para poder compararlas con la SNR de las imágenes clínicas. Como se mencionó anteriormente, los resultados indican que las imágenes de investigación tienen mayor nitidez que las imágenes de investigación simuladas, mientras que las imágenes clínicas (imágenes objetivo) resultan más nítidas que las imágenes predichas aunque la SNR de ambas es cercana, y al mismo tiempo se puede observar un cambio evidente en la SNR de las imágenes de investigación a la SNR de las imágenes predichas, indicando que el ruido está disminuyendo y la calidad de las imágenes de investigación está mejorando al procesarlas con el autocodificador.

	Imágenes de investigación	Investigación simuladas	Imágenes objetivo	Predichas
SNR	7.3453	6.9041	10.2885	9.6327

5.3. Segmentación de lesiones en imágenes de investigación

Como se explicó anteriormente, se utilizó la arquitectura STAN para realizar la segmentación en las imágenes de investigación. Primero se realizó la réplica de la red, la cual se puede encontrar en [7] con los mismos datos que utilizan en [33] para verificar si la construcción de la red era correcta. A continuación se presentan los resultados.

5.3.1. Arquitectura STAN con imágenes clínicas

En el artículo se utilizan dos colecciones de imágenes privadas, de las cuales sólo se pudo obtener acceso a una de ellas, por lo que todas las comparaciones se hicieron solamente con la llamada *dataset B* [41]. Esta colección de imágenes posee 163 imágenes clínicas, de las cuales, el 80 % se ocupó para entrenamiento, 10 % de validación y 10 % de prueba.

La red se entrenó por 50 épocas con un *learning rate* de 0.0001, utilizando el optimizador Adam. La función de pérdida utilizada es la función de pérdida Dice, la cual calcula el error de la superposición entre la segmentación predicha y la máscara original.

El comportamiento de la pérdida durante el entrenamiento se muestra en la siguiente gráfica:

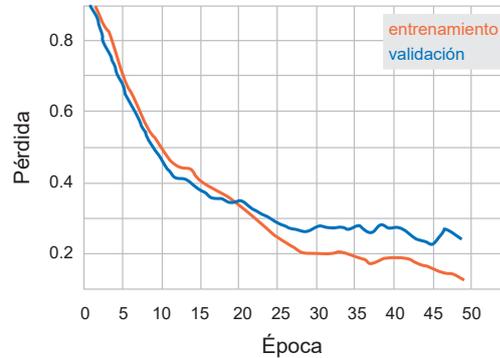


Figura 5.6: Comportamiento de la pérdida.

Para realizar la comparación entre la construcción de la red y el artículo, se calcularon 7 métricas: la tasa de verdaderos positivos (TPR), tasa de falsos positivos (FPR), tasa de verdaderos negativos (TNR), tasa de falsos negativos (FNR), índice Jaccard (JI), coeficiente de Dice (DSC) y la tasa de error del área (AER).

	TPR	FPR	TNR	FNR	JI	DSC	AER
Artículo	0.801	0.266	0.734	0.199	0.695	0.782	0.465
Réplica	0.749	0.004	0.995	0.250	0.663	0.783	0.431

A continuación se muestran algunos ejemplos de las segmentaciones predichas por esta red, teniendo en la primera columna la imagen original, en la segunda la máscara marcada por expertos, en la tercera la máscara predicha, en la cuarta la resta de la máscara original menos la máscara predicha, mostrando en rojo los falsos positivos y en verde los falsos negativos de la predicción, y en la quinta columna las máscaras sobrepuestas en la imagen original, mostrando en verde la máscara original y en rojo la máscara predicha:



Figura 5.7: Col 1: imagen original, col 2: máscara original, col 3: máscara predicha, col 4: resta de máscaras, col 5: máscaras sobrepuestas.

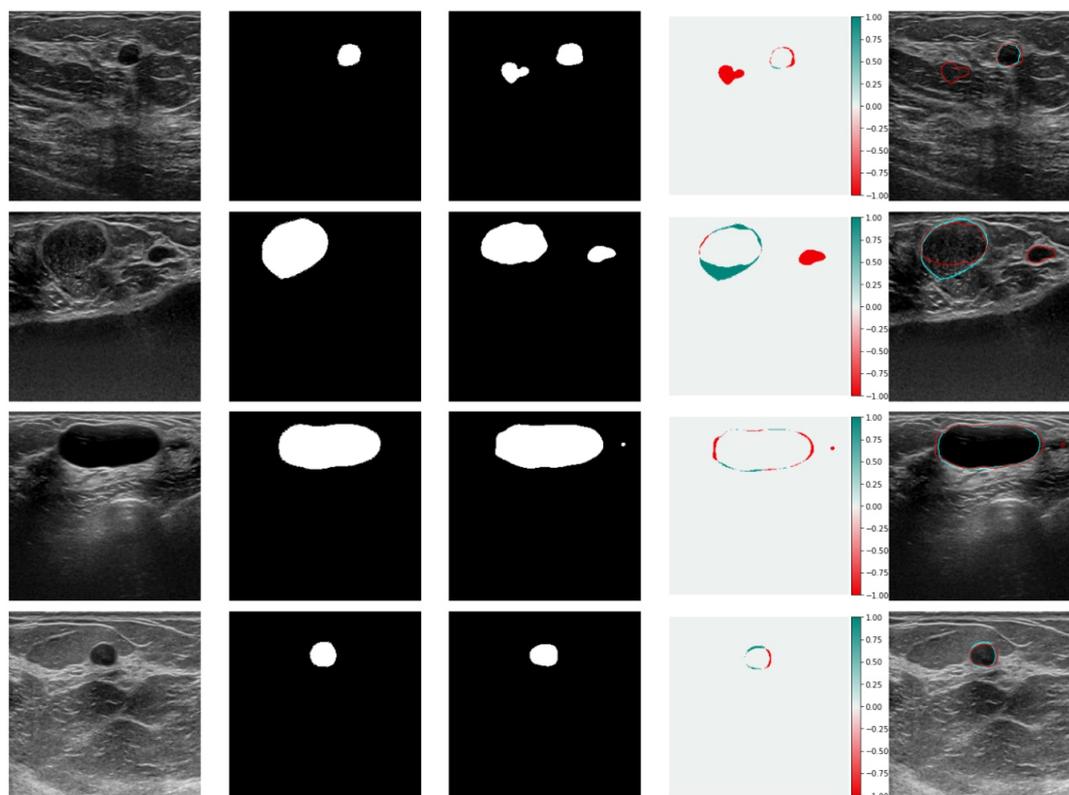


Figura 5.8: Col 1: imagen original, col 2: máscara original, col 3: máscara predicha, col 4: resta de máscaras, col 5: máscaras superpuestas.

Como se puede observar, el desempeño de la red es parecido a lo que dice el artículo, excepto que se tienen menos verdaderos positivos y más falsos negativos, pero en cuanto a la comparación de las áreas los resultados son bastante similares. Una vez que se comprobó que la red estaba arrojando los resultados que indican en el artículo se prosiguió con los siguientes pasos.

5.3.2. Arquitectura STAN con imágenes de investigación con reducción de patrón de moteado

Para esta prueba, la colección de imágenes que se ocupó fue de 71 imágenes procesadas con el autocodificador, a la cual se le realizó aumento de datos con la biblioteca *imgaug*, aplicando transformaciones de escalamiento, rotaciones de -45 a 45 grados y espejos horizontales (Figura 5.9), dando como resultado 3266 imágenes, de las cuales el 80 % se ocupó para entrenamiento, 10 % de validación y 10 % de prueba.

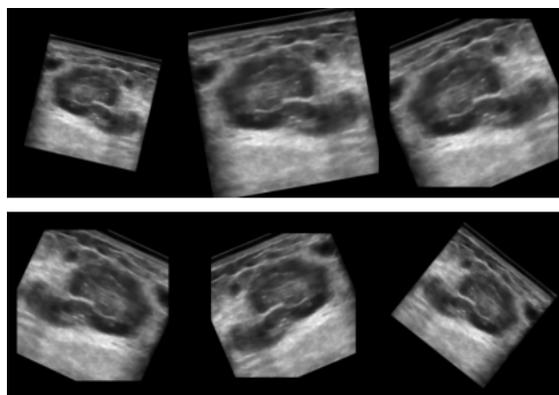


Figura 5.9: Imágenes después de aplicar aumento de datos.

La red se entrenó por 200 épocas con un *learning rate* de 0.00001, utilizando el optimizador Adam y la función de pérdida Dice. Por otro lado, se implementó el paro temprano de la red si la pérdida no disminuía durante 15 épocas para evitar el sobreajuste de ésta.

El comportamiento de la pérdida durante el entrenamiento se muestra en la siguiente gráfica:

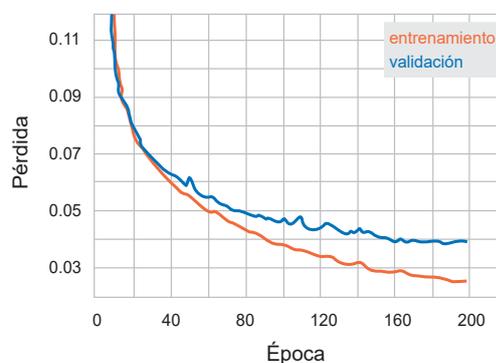


Figura 5.10: Comportamiento de la pérdida.

Para analizar el comportamiento de la red se utilizaron 10 métricas: la tasa de verdaderos positivos (TPR), tasa de falsos positivos (FPR), tasa de verdaderos negativos (TNR), tasa de falsos negativos (FNR), coeficiente de Dice (DSC), tasa de error del área (AER), sensibilidad, precisión, especificidad y exactitud. En la primera tabla se puede ver la matriz de confusión de los valores reales (especificados por las radiólogas) y los valores predichos por la red, mientras que en la segunda tabla se pueden ver los

5. RESULTADOS

resultados de las métricas restantes:

		Predicción	
		T	F
Real	T	0.93	0.06
	F	0.007	0.99

	DSC	AER	SENS.	ESPEC.	EXAC.	PREC.
Modelo autocodificador	0.91	0.16	0.93	0.99	0.96	0.99

A continuación se muestran algunos ejemplos de las segmentaciones predichas por esta red, teniendo en la primera columna la imagen original, en la segunda la máscara marcada por las expertas, en la tercera la máscara predicha, en la cuarta la resta de la máscara original menos la máscara predicha, mostrando en rojo los falsos positivos y en verde los falsos negativos de la predicción, y en la quinta columna las máscaras superpuestas en la imagen original, mostrando en verde la máscara original y en rojo la máscara predicha:

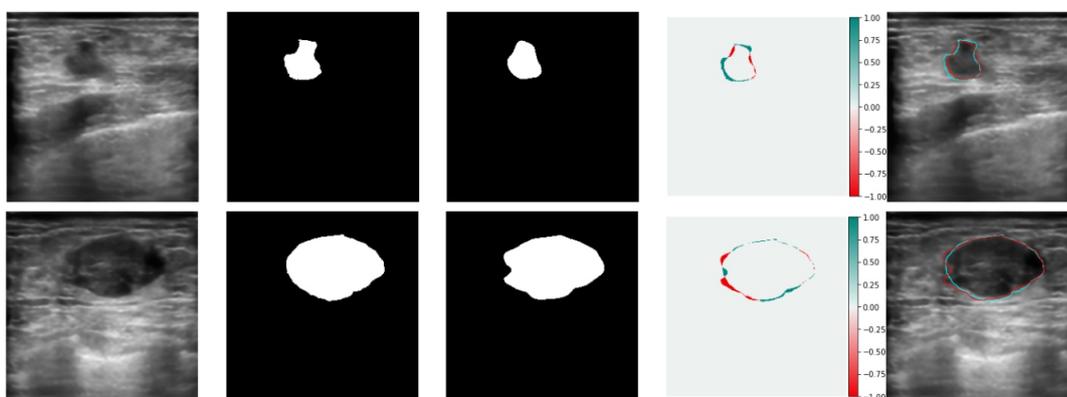


Figura 5.11: Col 1: imagen original, col 2: máscara original, col 3: máscara predicha, col 4: resta de máscaras, col 5: máscaras superpuestas.

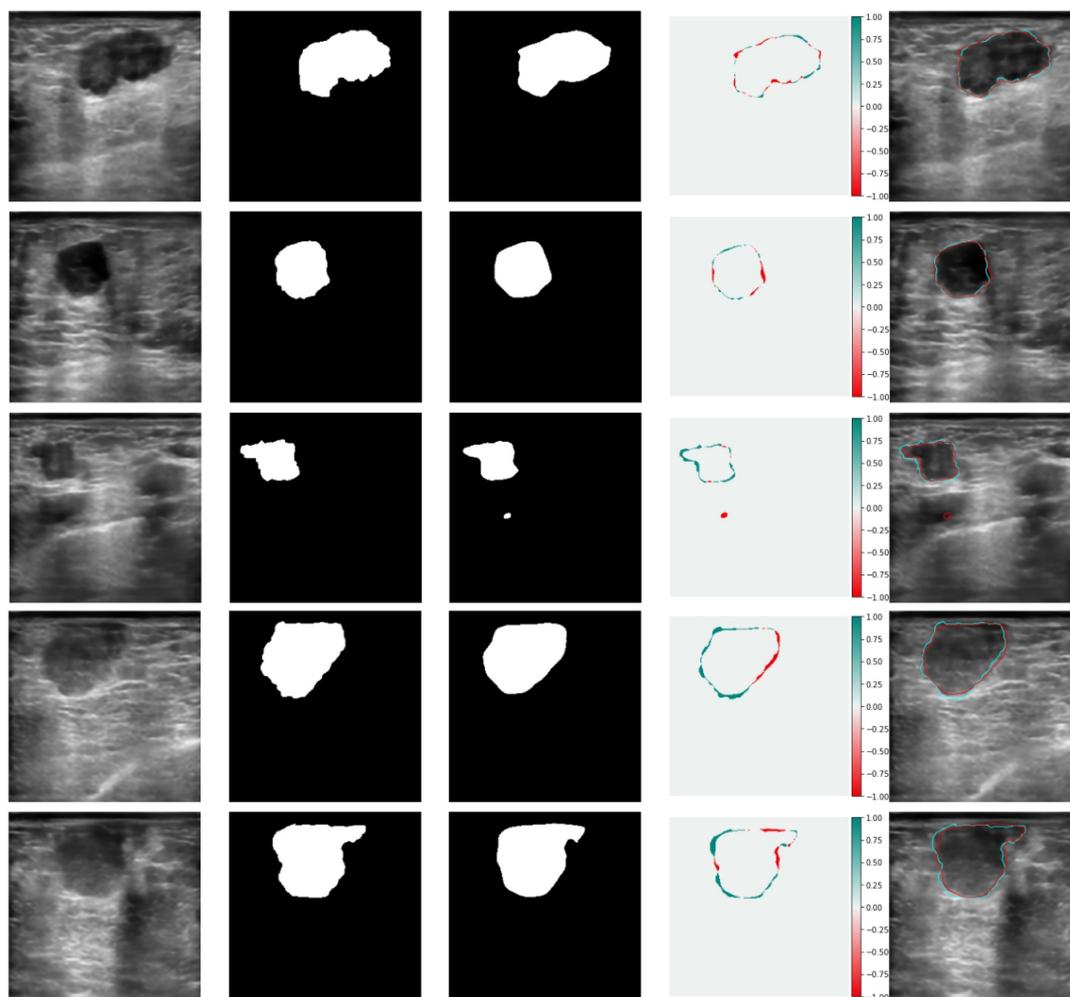


Figura 5.12: Col 1: imagen original, col 2: máscara original, col 3: máscara predicha, col 4: resta de máscaras, col 5: máscaras sobrepuestas.

Como se puede observar, este modelo arroja resultados bastante altos gracias a la limpieza que se realizó con el autocodificador y al aumento de datos utilizado.

El índice Dice indica que las predicciones están siendo 91 % parecidas a las máscaras marcadas por las radiólogas, mientras que la tasa de error del área está siendo muy pequeña, lo que indica que no hay mucha presencia de falsos positivos en las máscaras predichas.

5.3.3. Arquitectura STAN con dos *datasets* realizando inicialización de los pesos

Esta prueba se realizó con la intención de tener un sistema de segmentación más robusto debido a que el entrenamiento se realiza con dos *datasets* de distinto tipo: un *dataset* contiene mucho ruido y otro *dataset* contiene imágenes limpias.

La colección de imágenes que se ocupó fue de 142 imágenes: 71 imágenes de investigación y 71 imágenes procesadas con el autocodificador; de igual forma se realizó aumento de datos con la biblioteca *imgaug*, aplicando transformaciones de escalamiento, rotaciones de -45 a 45 grados y espejos horizontales, dando como resultado 3550 imágenes, de las cuales el 80 % se ocupó para entrenamiento, 10 % de validación y 10 % de prueba.

La red se inicializó con los pesos del modelo previamente entrenado y se decrementó el *learning rate* a 0.000001 entrenando por 200 épocas, de igual forma se utilizó el optimizador Adam y la función de pérdida Dice. Además, se implementó el paro temprano de la red si la pérdida no disminuía durante 15 épocas para evitar el sobreajuste de ésta.

El comportamiento de la pérdida durante el entrenamiento se muestra en la siguiente gráfica:

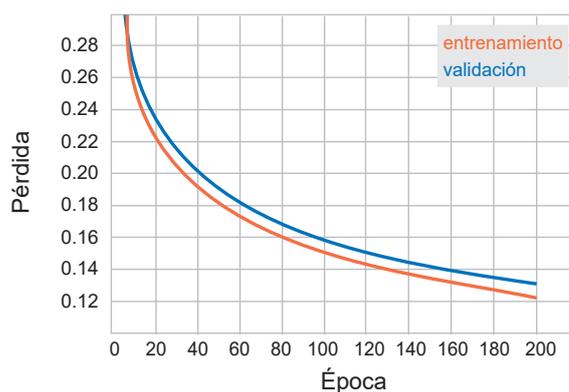


Figura 5.13: Comportamiento de la pérdida.

Para analizar el comportamiento de la red se utilizaron 10 métricas: la tasa de verdaderos positivos (TPR), tasa de falsos positivos (FPR), tasa de verdaderos negativos

(TNR), tasa de falsos negativos (FNR), coeficiente de Dice (DSC), tasa de error del área (AER), sensibilidad, precisión, especificidad y exactitud. En la primera tabla se puede ver la matriz de confusión de los valores reales (especificados por las radiólogas) y los valores predichos por la red, mientras que en la segunda tabla se pueden ver los resultados de las métricas restantes:

		Predicción	
		T	F
Real	T	0.94	0.05
	F	0.01	0.98

	DSC	AER	SENS.	ESPEC.	EXAC.	PREC.
Modelo robusto	0.91	0.15	0.94	0.98	0.96	0.98

A continuación se muestran algunos ejemplos de las segmentaciones predichas por esta red, teniendo en la primera columna la imagen original, en la segunda la máscara marcada por las expertas, en la tercera la máscara predicha, en la cuarta la resta de la máscara original menos la máscara predicha, mostrando en rojo los falsos positivos y en verde los falsos negativos de la predicción, y en la quinta columna las máscaras sobrepuestas en la imagen original, mostrando en verde la máscara original y en rojo la máscara predicha:



Figura 5.14: Col 1: imagen original, col 2: máscara original, col 3: máscara predicha, col 4: resta de máscaras, col 5: máscaras sobrepuestas.

5. RESULTADOS

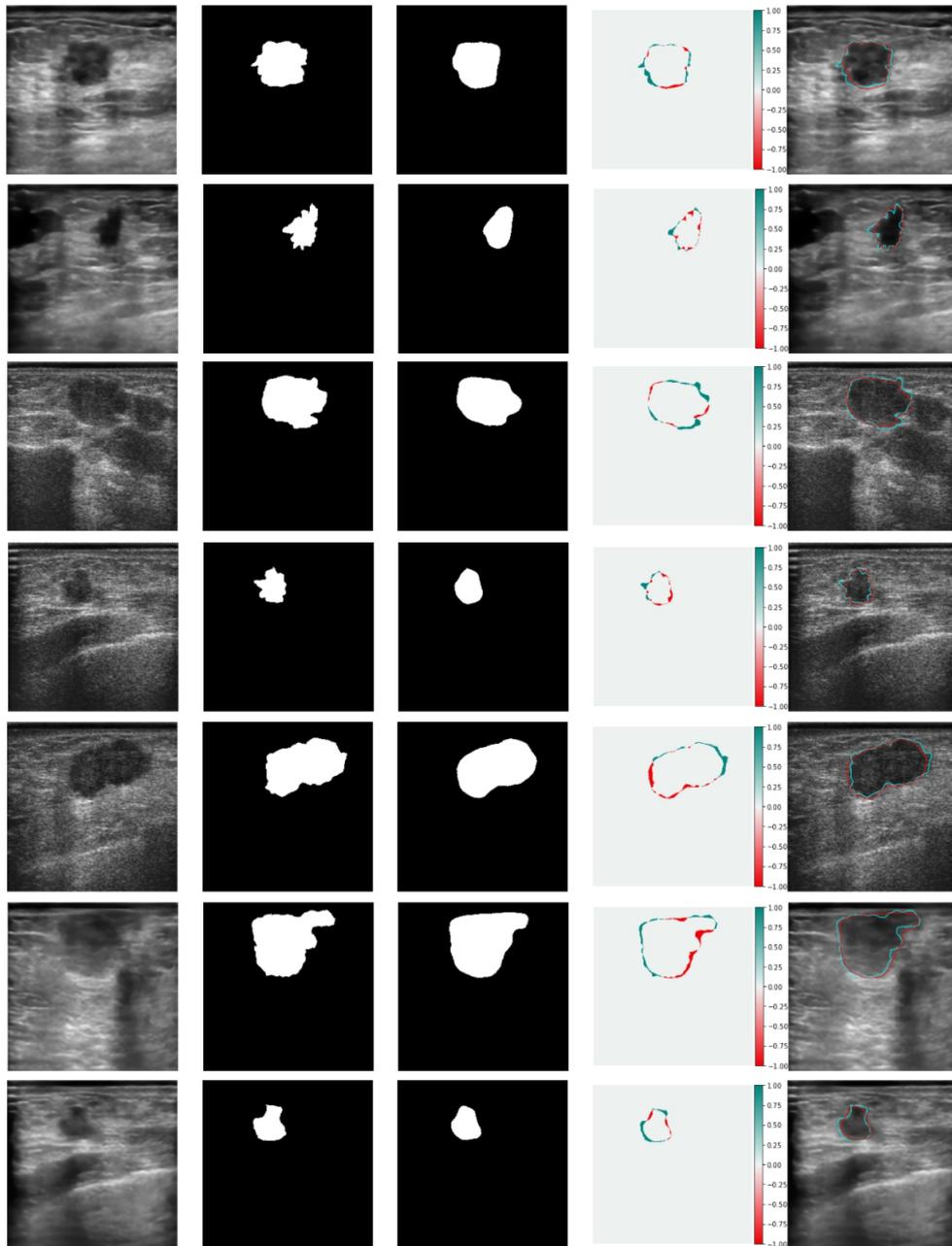


Figura 5.15: Col 1: imagen original, col 2: máscara original, col 3: máscara predicha, col 4: resta de máscaras, col 5: máscaras sobrepuestas.

El índice Dice indica que las predicciones también están siendo 91% parecidas a las máscaras marcadas por las radiólogas, mientras que la tasa de error del área está bajando a 0.15, lo que indica que hay menos presencia de falsos positivos en las máscaras

predichas.

Como se puede observar, este modelo arroja resultados bastante altos y muy parecidos al modelo anterior, lo que indica que el modelo es más robusto ya que también es capaz de detectar lesiones en las imágenes de investigación.

Sin embargo, que se obtengan resultados altos en el modelo no quiere decir que realmente sea bueno para la tarea que desempeñará; para poder indicar que tiene un buen desempeño realmente, se debe comparar contra el error de segmentación que existe entre los radiólogos, de esta manera se podrá saber si el modelo está realizando predicciones parecidas a las que podría hacer otro radiólogo.

5.3.4. Cálculo de errores entre radiólogas

Como ya se ha mencionado, se cuenta con una colección de imágenes de 71 imágenes de investigación, cada una marcada manualmente por dos radiólogas del INCan (Figura 5.16); con estas dos referencias es posible obtener el error que existe entre ellas para poder comparar el desempeño de las segmentaciones manuales con los modelos de redes neuronales.



Figura 5.16: Izq: Segmentación manual de la radióloga 1, der: segmentación manual de la radióloga 2.

5. RESULTADOS

Las métricas que se obtuvieron son las mismas 10 que en los modelos: la tasa de verdaderos positivos (TPR), tasa de falsos positivos (FPR), tasa de verdaderos negativos (TNR), tasa de falsos negativos (FNR), coeficiente de Dice (DSC), tasa de error del área (AER), sensibilidad, precisión, especificidad y exactitud.

La manera en que se realizaron las mediciones fue tomar como referencia a la radióloga 1 y obtener las medidas con respecto a la radióloga 2, y posteriormente tomar como referencia a la radióloga 2 y obtener las medidas con respecto a la radióloga 1. Estas mediciones se promediaron para obtener los resultados que se ocuparán para comparar contra los modelos de las redes.

Las siguientes tablas muestran los promedios de las métricas:

		Predicción	
		T	F
Real	T	0.92	0.07
	F	0.01	0.98

	DSC	AER	SENS.	ESPEC.	EXAC.	PREC.
Radiólogas	0.92	0.15	0.92	0.98	0.95	0.98

5.3.5. Comparación del desempeño de los distintos modelos con el error entre radiólogas

Finalmente se tiene la comparación entre radiólogo y predicciones de la red:

	TPR	FPR	TNR	FNR
Modelo autocodificador	0.93	0.007	0.99	0.06
Modelo robusto	0.94	0.01	0.98	0.05
Radiólogas	0.92	0.01	0.98	0.07

	DSC	AER	SENS.	ESPEC.	EXAC.	PREC.
Modelo autocodificador	0.91	0.16	0.93	0.99	0.96	0.99
Modelo robusto	0.91	0.15	0.94	0.98	0.96	0.98
Radiólogas	0.92	0.15	0.92	0.98	0.95	0.98

Como se puede ver, los resultados de los dos modelos entrenados y las radiólogas son muy parecidos, superando incluso en las métricas de sensibilidad y exactitud a las radiólogas. Es por eso que se puede afirmar que la red sí está teniendo buen desempeño con este conjunto de datos, de tal forma que sus predicciones podrían compararse con las segmentaciones que pueda hacer otro radiólogo.

5.4. Discusión de resultados

En el caso del autocodificador se puede apreciar a simple vista que la imagen reconstruida sí está disminuyendo el patrón de moteado en gran medida, además que el SSIM indica que se está conservando la estructura de la mama, lo que era muy importante al realizar esta tarea ya que los radiólogos deben ser capaces de ver la estructura de la mama para realizar una detección y segmentación exitosas. Por otro lado, que aumente la SNR indica que la calidad de la imagen predicha es mejor con respecto a la imagen de investigación que se utilizó en la entrada y, al mismo tiempo, que sea cercana a la SNR de las imágenes clínicas originales indica que tiene una calidad parecida a éstas.

5. RESULTADOS

En el caso de los modelos entrenados con la arquitectura STAN, se puede afirmar que a pesar de que el artículo no es lo suficientemente claro con la arquitectura y existen algunas ambigüedades como que no menciona qué funciones de activación se ocupan (salvo en la capa de salida), hay errores entre lo que dice el gráfico de la arquitectura contra lo que se menciona en el artículo, hay discrepancias en las definiciones del codificador y decodificador, no se menciona cuántos filtros se aplican en cada capa, etc., se lograron obtener resultados bastante similares al realizar la réplica de la red.

Por otro lado, en el modelo entrenado con las imágenes procesadas por el autocodificador, se está haciendo un excelente trabajo reconociendo lesiones de distinto tamaño y la red es capaz de detectar muchos detalles gracias a los diferentes tamaños de ventanas que se están utilizando en los filtros, así como las múltiples conexiones de salto, además de la modificación del número de épocas de entrenamiento y del *learning rate*, el cual se hizo más pequeño para mejorar el aprendizaje de la red. Del mismo modo, inicializar los pesos y sesgos en el segundo modelo con el modelo entrenado con las imágenes del autocodificador y bajar más el *learning rate*, ayuda a tener muy buenos resultados en las predicciones y a tener un modelo más robusto ya que es capaz de detectar lesiones tanto en imágenes clínicas como de investigación.

El contar con las segmentaciones realizadas por dos radiólogas da la oportunidad de saber si el modelo es tan bueno para predecir las lesiones como si lo hiciera otro radiólogo, para que de esta forma se pueda asegurar que se está teniendo buen desempeño de la red o saber si se tienen que hacer modificaciones. En este caso, las predicciones tanto del modelo entrenado con imágenes procesadas con el autocodificador como del modelo robusto son bastante parecidas a las que podría realizar otro experto, por lo que ahora sí se puede asegurar que el desempeño de estos modelos es lo suficientemente bueno para detectar lesiones en este conjunto de datos.

Finalmente, se deben tener claras las limitaciones de esta tesis. Por un lado se cuenta con muy pocas imágenes segmentadas manualmente, y por el otro solamente se tienen los trazos de dos radiólogas, por lo que los resultados de este trabajo están sujetos a esas dos condiciones. Para poder asegurar que las segmentaciones predichas por los modelos se parecen a los trazos que podría hacer otro radiólogo, se deben recabar más segmentaciones manuales realizadas por un número más grande de expertos para

tener un error de segmentación más exacto y poder comparar contra las métricas de los modelos.

Por otro lado, para mejorar los modelos se debe tener una colección de imágenes mucho más grande, ya que, aunque realizar aumento de datos puede resultar en *data-sets* mucho mayores, aplicarlo en dominios tan limitados podría resultar en un mayor sobreajuste [34]. Es por lo anterior que se deben recabar imágenes de más pacientes y tipos de lesiones para que los modelos puedan aprender a identificar mejor cualquier tipo de lesión, ya sea maligna o benigna, y de cualquier tamaño.

Conclusiones

Se propuso una nueva forma de recabar las máscaras manuales haciendo que los radiólogos puedan tener acceso a una aplicación en un dispositivo móvil al que sólo se le tendrían que cargar las imágenes por segmentar. De esta forma se logra que el médico tenga una buena experiencia al realizar esta tarea y deje de ser tardada y tediosa. Se espera que la aplicación pueda estar disponible para cualquier médico o persona que busque realizar tareas de este tipo.

De igual manera, se logró diseñar el algoritmo que permite obtener las máscaras de segmentación basado en los contornos de referencia obtenidos desde la aplicación. Este algoritmo se encuentra disponible en los enlaces de github proporcionados anteriormente.

También se logró diseñar e implementar un algoritmo que ayuda a reducir el patrón de moteado de las imágenes de investigación, esto para poder tener imágenes parecidas a las clínicas, a las cuales se les aplican filtros que ayudan a que la imagen se vea muy definida. La arquitectura utilizada para este algoritmo se basa en la arquitectura U-Net pero se le eliminan las conexiones de salto para que se tenga una arquitectura pura de autocodificador. Los resultados que entrega esta red indican que se conserva la estructura de la mama y que las imágenes predichas son 91 % parecidas a las imágenes clínicas. Además, la SNR indica que la calidad de las imágenes aumentó al procesarlas con el autocodificador, lo que quiere decir que se está disminuyendo el patrón de moteado de las imágenes.

Por otro lado, aunque el artículo no describe de la mejor forma la arquitectura de STAN y se tienen muchas ambigüedades, se logró replicar exitosamente la red,

6. CONCLUSIONES

obteniendo resultados muy parecidos a los publicados con el *dataset* B.

Una vez que se comprobó que la red estaba arrojando los resultados que se indican en el artículo, se le modificaron parámetros de compilación y entrenamiento importantes: el *learning rate* y el número de épocas en las que entrena. Se entrenó con las imágenes de investigación procesadas con el autocodificador y se hizo uso de la biblioteca de Python *imgaug* para realizar aumento de datos. Esas modificaciones ayudaron a que la red aprendiera a detectar mejor las características de las imágenes, obteniendo un índice Dice de 0.91 con una tasa de error de área de 0.16, lo que indica que las predicciones están siendo 91 % parecidas a las máscaras realizadas por las expertas. De igual manera, este algoritmo se encuentra disponible en los enlaces de github proporcionados.

El modelo anterior se utilizó en un tercer experimento para poder inicializar los pesos de la red, la cual se entrenó con un *dataset* más grande debido a que se utilizaron las imágenes de investigación así como las imágenes procesadas con el autocodificador, además se hizo uso de la biblioteca de Python *imgaug* para realizar aumento de datos y también se modificó el *learning rate*. Se obtienen resultados de índice Dice de 0.91 y tasa de error de área de 0.15, y aunque no hayan mejorado drásticamente los resultados con respecto al modelo anterior, lo que se logra con esto es tener un modelo de segmentación automática más robusto ya que es capaz de detectar lesiones en distintos tipos de imagen.

Aunque los resultados de los modelos son altos en el índice Dice y bajos en la tasa de error de área, para asegurar que la segmentación podría asemejarse a la de un experto se tiene que validar el método realizando la comparación entre los resultados del modelo con el error que exista entre radiólogos. Debido a que se tienen dos segmentaciones manuales para cada imagen realizada por dos radiólogas es que se puede hacer esta comparación.

Además, como se puede ver en los resultados, las predicciones de los modelos tienen métricas bastante similares a las que se tienen con las segmentaciones manuales, ya que las radiólogas tienen un índice Dice de 0.92 y una tasa de error de área de 0.15, por lo que ahora sí se puede asegurar que el modelo tiene muy buen desempeño, ya que es capaz de predecir las máscaras como si de otro experto se tratara, logrando así que se cumpla el objetivo de este trabajo.

Finalmente, es importante reconocer los alcances y limitaciones de esta tesis que son el número de imágenes de investigación con las que se cuenta y el número de radiólogos que realizan la segmentación manual. Estos son puntos que continuarán trabajándose ya que se espera que a futuro se desarrolle un sistema de segmentación en línea para poder ayudar a los médicos a ser más exactos en los diagnósticos.

6.1. Trabajo futuro

Se desea conseguir que el sistema de segmentación esté disponible en una plataforma web, el cual tenga como entrada una imagen de ultrasonido de mama sin marcar y como salida una imagen filtrada y segmentada. Esta plataforma estará alojada en el servidor del Instituto de Física de la UNAM¹ y podrá ser accesible para cualquier persona que se registre.

La razón de realizar un sistema de segmentación en línea es que existen hospitales con equipos de ultrasonido muy avanzados que incluso estos mismos pueden hacer un filtrado de la imagen muy bueno y segmentar las imágenes, pero al mismo tiempo hay hospitales que siguen contando con equipos que entregan imágenes de muy mala calidad, lo que es un problema para los médicos ya que tardan más en hacer el marcado de la imagen y se requiere mayor experiencia de su parte, además de reducir la reproducibilidad de la interpretación de las imágenes de ultrasonido. Tener este sitio disponible puede ayudarles a ser más exactos en los diagnósticos y les puede ahorrar tiempo al darles una idea aproximada de la región con alguna lesión.

Al realizar la búsqueda de información de segmentación de imágenes no se encontró un sistema en México que pueda ayudar a segmentar las imágenes de ultrasonido de mama en línea, dando la oportunidad de ser los primeros en tener un sitio destinado a la segmentación. Esto contribuirá a la transferencia social del conocimiento y la tecnología desarrollados durante esta tesis.

¹Estará alojada en el servidor una vez que se tenga el permiso del Instituto de Física de la UNAM.

Bibliografía

- [1] (2021). Breast biopsy. <https://www.mayoclinic.org/tests-procedures/breast-biopsy/about/pac-20384812>. 31-07-2021. 15
- [2] (2021). The challenge of dense breasts. <http://www.metritrack.com/the-challenge-of-dense-breasts>. 21-04-2021. 11
- [3] (2021). Cáncer de mama. <https://www.who.int/es/news-room/fact-sheets/detail/breast-cancer>. 26-03-2021. 7, 8
- [4] (2021). Dense breast tissue: What it means to have dense breasts. <https://www.mayoclinic.org/tests-procedures/mammogram/in-depth/dense-breast-tissue/art-20123968>. 22-04-2021. 2
- [5] Al-Dhabyani, W., Gomaa, M., Khaled, H., and Fahmy, A. (2020). Dataset of breast ultrasound images. *Data in brief*, 28:104863. 38
- [6] Baraldi, A. and Parmiggiani, F. (1995). A refined gamma map sar speckle filter with improved geometrical adaptivity. *IEEE Transactions on Geoscience and Remote Sensing*, 33(5):1245–1257. 17
- [7] Bass, V. (2021). Segmentación automática de lesiones en imágenes de ultrasonido de mama utilizando redes neuronales convolucionales: códigos. <https://github.com/VivianBass/Tesis>. 51, 56
- [8] Brandan, M. E. and Villaseñor, Y. (2006). Detección del cáncer de mama: estado de la mamografía en México. *Cancerología*, 1(3):147–62. 9
- [9] Cárdenas-Sánchez, J., Bargalló-Rocha, E., Erazo-Valle, A., Chacón, A., Valero-Castillo, V., and Pérez-Sánchez, V. (2019). Consenso mexicano sobre diagnóstico

BIBLIOGRAFÍA

- y tratamiento del cáncer mamario. octava revisión. colima 2019. *Gaceta Mexicana de Oncología*, 18(3):1. 1
- [10] Chan, V. and Perlas, A. (2011). Basics of ultrasound imaging. In *Atlas of ultrasound-guided procedures in interventional pain management*, pages 13–19. Springer. 10
- [11] Cheng, H.-D., Shan, J., Ju, W., Guo, Y., and Zhang, L. (2010). Automated breast cancer detection and classification using ultrasound images: A survey. *Pattern recognition*, 43(1):299–317. 16
- [12] Equipo de redactores y equipo de editores médicos de la Sociedad Americana Contra El Cáncer” (2019). Imágenes por resonancia magnética (MRI) de los senos. <https://www.cancer.org/es/cancer/cancer-de-seno/pruebas-de-deteccion-y-deteccion-temprana-del-cancer-de-seno/imagenes-por-resonancia-magnetica-de-los-senos.html>. 21-04-2021. 14
- [13] Fernández, J. Á., Ozores, P. P., López, V. C., Mosquera, A. C., and López, R. L. (2021). Cáncer de mama. *Medicine-Programa de Formación Médica Continuada Acreditado*, 13(27):1506–1517. 7
- [14] Frost, V. S., Stiles, J. A., Shanmugan, K. S., and Holtzman, J. C. (1982). A model for radar images and its application to adaptive digital filtering of multiplicative noise. *IEEE Transactions on pattern analysis and machine intelligence*, (2):157–166. 17
- [15] Harwood, D., Subbarao, M., Hakalahti, H., and Davis, L. S. (1987). A new class of edge-preserving smoothing filters. *Pattern Recognition Letters*, 6(3):155–162. 18
- [16] He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778. 29
- [17] Hongyu and Wendy (2019). Recitation – Soft k-means clustering. <https://www.cs.cmu.edu/~02251/recitations/recitation-soft-clustering.pdf>. 20-10-2020. 23

- [18] Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q. (2017a). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708. [29](#)
- [19] Huang, Q., Luo, Y., and Zhang, Q. (2017b). Breast ultrasound image segmentation: a survey. *International journal of computer assisted radiology and surgery*, 12(3):493–507. [25](#)
- [20] Isensee, F., Petersen, J., Klein, A., Zimmerer, D., Jaeger, P. F., Kohl, S., Wasserthal, J., Koehler, G., Norajitra, T., Wirkert, S., et al. (2018). nnu-net: Self-adapting framework for u-net-based medical image segmentation. *arXiv preprint arXiv:1809.10486*. [43](#)
- [21] Jensen, J. A. (1996). Field: A program for simulating ultrasound systems. In *10TH NORDICBALTIC CONFERENCE ON BIOMEDICAL IMAGING, VOL. 4, SUPPLEMENT 1, PART 1: 351–353*. Citeseer. [38](#), [39](#)
- [22] Knaul, F. M., Nigenda, G., Lozano, R., Arreola-Ornelas, H., Langer, A., and Frenk, J. (2009). Cáncer de mama en México: una prioridad apremiante. *Salud pública de México*, 51:s335–s344. [1](#)
- [23] Kockanat, S. and Karaboga, N. (2015). A novel 2d-abc adaptive filter algorithm: a comparative study. *Digital Signal Processing*, 40:140–153. [13](#)
- [24] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105. [29](#)
- [25] Kuan, D. T., Sawchuk, A. A., Strand, T. C., and Chavel, P. (1985). Adaptive noise smoothing filter for images with signal-dependent noise. *IEEE transactions on pattern analysis and machine intelligence*, (2):165–177. [17](#)
- [26] Lee, J.-S. (1980). Digital image enhancement and noise filtering by use of local statistics. *IEEE transactions on pattern analysis and machine intelligence*, (2):165–168. [17](#)
- [27] Pérez Zúñiga, I., Villaseñor Navarro, Y., Pérez Badillo, M. P., Cruz Morales, R., Pavón Hernández, C., and Aguilar Cortázar, L. O. (2012). Resonancia magnética de mama y sus aplicaciones. *Gaceta mexicana de oncología*, 11(4):268–280. [14](#)

- [28] Pickens, D. R., Li, Y., Morgan, V. L., and Dawant, B. M. (2005). Development of computer-generated phantoms for fmri software evaluation. *Magnetic resonance imaging*, 23(5):653–663. [39](#)
- [29] Rafati, M., Arabfard, M., Zadeh, M. R. R., and Maghsoudloo, M. (2016). Assessment of noise reduction in ultrasound images of common carotid and brachial arteries. *IET Computer Vision*, 10(1):1–8. [18](#)
- [30] Robles, F. T. (2017). *Adquisición y análisis automático de imágenes de ultrasonido de tumores de mama*. [13](#)
- [31] Rodríguez López, V. et al. (2012). Análisis de imágenes de mamografía para la detección de cáncer de mama. *REPOSITORIO NACIONAL CONACYT*. [9](#)
- [32] Shan, J., Cheng, H., and Wang, Y. (2012). A novel segmentation method for breast ultrasound images based on neutrosophic l-means clustering. *Medical physics*, 39(9):5669–5682. [23](#)
- [33] Shareef, B., Xian, M., and Vakanski, A. (2020). Stan: Small tumor-aware network for breast ultrasound image segmentation. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pages 1–5. IEEE. [VIII](#), [4](#), [5](#), [43](#), [48](#), [56](#)
- [34] Shorten, C. and Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of big data*, 6(1):1–48. [69](#)
- [35] Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*. [29](#)
- [36] Stavros, A. T., Thickman, D., Rapp, C. L., Dennis, M. A., Parker, S. H., and Sisney, G. A. (1995). Solid breast nodules: use of sonography to distinguish between benign and malignant lesions. *Radiology*, 196(1):123–134. [2](#)
- [37] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9. [29](#)

- [38] Thigpen, D., Kappler, A., and Brem, R. (2018). The role of ultrasound in screening dense breasts—a review of the literature and practical solutions for implementation. *Diagnostics*, 8(1):20. [10](#)
- [39] Tomasi, C. and Manduchi, R. (1998). Bilateral filtering for gray and color images. In *Sixth international conference on computer vision (IEEE Cat. No. 98CH36271)*, pages 839–846. IEEE. [18](#)
- [40] Yang, J., Fan, J., Ai, D., Wang, X., Zheng, Y., Tang, S., and Wang, Y. (2016). Local statistics and non-local mean filter for speckle noise reduction in medical ultrasound image. *Neurocomputing*, 195:88–95. [13](#)
- [41] Yap, M. H., Pons, G., Marti, J., Ganau, S., Sentis, M., Zwiggelaar, R., Davison, A. K., and Marti, R. (2017). Automated breast ultrasound lesions detection using convolutional neural networks. *IEEE journal of biomedical and health informatics*, 22(4):1218–1226. [1](#), [56](#)
- [42] Yu, Y. and Acton, S. T. (2002). Speckle reducing anisotropic diffusion. *IEEE Transactions on image processing*, 11(11):1260–1270. [18](#)