

Diseño e Implementación de un Algoritmo para la Estimación y Corrección de la Deformación
del Seno en Mamografías Digitales mediante Procesamiento de Imágenes

Luis Alejandro Pedreros Martínez y Paula Andrea Mora Escudero

Trabajo de Grado para Optar el Título de Ingeniero Electrónico

Directo

Said David Pertuz Arroyo

Doctor en Ingeniería Electrónica

Co-Directora

Mónica Andrea Botero Londoño

Doctora en Ingeniería Electrónica

Universidad Industrial de Santander

Facultad de Ingenierías Físico-Mecánicas

Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones

Ingeniería Electrónica

Bucaramanga

2024

Dedicatoria

A mis padres por haberme orientado en todo este camino con su amor y apoyo incondicional. A mis abuelos por su cariño, compañía y las sabias enseñanzas que día a día me han brindado. A mi abuelo Ismael, pues llevaré conmigo siempre su memoria y parte de la esencia de mi ser. A mis hermanos porque son la alegría que la vida me brindó para seguir adelante. A Alejandra, el amor con quien tuve la suerte de coincidir y crecer de la mano. A los demás familiares y amigos que hacen parte de ese logro.

Luis Alejandro Pedreros Martínez

A mi familia, fuente inagotable de amor y apoyo. A mi nona y mi tía, que con su cariño y sabiduría han estado siempre a mi lado. A mis padres y hermanos, por confiar en mi potencial y en lo que puedo lograr, incluso en los momentos de duda. A mis amigos, que han sido el pilar que me ha sostenido en los momentos más difíciles, gracias por cada gesto y palabra de aliento que me han hecho sentir que nunca estoy sola. A "Identidad Estudiantil", gracias por ser el espacio donde he crecido y aprendido tanto, por las manos extendidas en momentos difíciles y por hacerme sentir parte de algo más grande y significativo. Con todo mi corazón, este logro es tanto suyo como mío.

Paula Andrea Mora Escudero

Agradecimientos

Nos gustaría expresar nuestra más profunda gratitud al Grupo de Investigación en Conectividad y Procesamiento de Señales (CPS) por permitirnos trabajar junto a ellos en este proyecto. Su constante apoyo y disposición para compartir su vasta experiencia fueron esenciales para la realización de nuestro trabajo. Estamos profundamente agradecidos por su orientación y el ambiente de trabajo enriquecedor que nos brindaron, lo que nos permitió enfrentar los desafíos de esta investigación con confianza y determinación. Su compromiso con la excelencia y su espíritu colaborativo fueron una fuente constante de inspiración.

Queremos extender un agradecimiento muy especial al Profesor PhD Said David Pertuz Arroyo, director de este proyecto. Su inquebrantable apoyo y guía fueron cruciales en los momentos en que nos sentimos perdidos. Su capacidad para motivarnos, proporcionarnos dirección y alentarnos a seguir adelante, incluso en los tiempos más desafiantes, fue invaluable. Gracias por creer en nosotros, por darnos ese empujón cuando más lo necesitábamos, y por estar siempre dispuesto a compartir su tiempo y conocimiento. Su confianza y dedicación no solo nos permitieron completar este trabajo, sino que también nos ayudaron a crecer como profesionales y como individuos.

Finalmente, queremos agradecer a nuestras familias y amigos, cuyo constante apoyo fue esencial para mantenernos enfocados y motivados a lo largo de este viaje. Su paciencia y comprensión durante las largas horas de trabajo hicieron posible que lográramos nuestros objetivos.

Contenido

Introducción	13
1 Antecedentes	17
1.1 Conceptos Clave	17
1.1.1 <i>Mamografía Digital</i>	18
1.1.2 <i>Compresión Mamaria</i>	18
1.1.3 <i>Inhomogeneidades de Intensidad</i>	19
1.1.4 <i>Histograma de una Imagen</i>	19
1.1.5 <i>Índice de Similitud Estructural (SSIM)</i>	19
1.1.6 <i>Error Cuadrático Medio (RMSE)</i>	20
1.2 Revisión de Estudios Previos	20
1.2.1 <i>Modelos de Parámetros Globales</i>	20
1.2.2 <i>Algoritmos de Detección de Bordes</i>	21
1.3 Métodos Actuales para la Corrección de Deformaciones	21
1.3.1 <i>Métodos Basados en Modelos de Parámetros Globales</i>	21
1.3.2 <i>Algoritmos de Detección de Bordes</i>	22
1.4 Justificación y Ventajas del Enfoque Propuesto	22
2 Materiales y Métodos	24

2.1	Base de Datos	24
2.1.1	<i>Generación de Phantoms y Mamografías</i>	25
2.1.2	<i>Categorías de Densidad Mamaria</i>	26
2.1.3	<i>Calidad de los Datos y Preprocesamiento</i>	27
2.2	Generación de Verdades Fundamentales (GT)	27
2.2.1	<i>Almacenamiento de Rebanadas del Phantom</i>	28
2.2.2	<i>Generación del Campo de Deformación 2D</i>	29
2.2.3	<i>Segmentación del Air Gap</i>	29
2.2.4	<i>Creación de la Máscara GT</i>	29
2.3	Corrección de Intensidad	30
2.3.1	<i>Preprocesamiento</i>	31
2.3.2	<i>Detección del Air Gap</i>	31
2.3.3	<i>Aplicación de la Corrección de Intensidad</i>	34
2.4	Evaluación de la Corrección de Intensidad	35
3	Experimentos y Resultados	37
3.1	Detección del Air Gap	37
3.2	Corrección de Intensidad	38
3.3	Análisis de Resultados	39
4	Conclusiones	42

5 Recomendaciones	44
--------------------------	-----------

Referencias	44
--------------------	-----------

Apéndice	48
-----------------	-----------

Lista de Figuras

Figura 1	Efecto de la deformación mamaria en las intensidades de imagen	14
Figura 2	El <i>Air gap</i> es el espacio vacío entre las paletas de compresión y el seno, y es la causa de variaciones de intensidad en la mamografía	26
Figura 3	Generación de <i>Ground Truth</i>	28
Figura 4	Pre-procesamiento y detección del <i>Air gap</i>	31
Figura 5	Histograma de la imagen mamográfica y su curva ajustada	32
Figura 6	Rendimiento de la corrección de intensidad	40

Lista de Tablas

Tabla 1	<i>Distancia promedio en píxeles en el Air gap detectado</i>	38
Tabla 2	<i>Calidad de la corrección de intensidad usando SSIM</i>	38
Tabla 3	<i>Calidad de la corrección de intensidad usando RMSE</i>	39

Lista de Apéndices

	pág.
Apéndice 1. Base de Datos	48
Apéndice 2. Algoritmo Propuesto	50

Glosario

BIRADS (Sistema de Informes y Datos de Imágenes de Mama): sistema estandarizado para informar y categorizar hallazgos mamográficos, que clasifica los senos en diferentes categorías según la densidad del tejido.

VICTRE (Ensayo Clínico Virtual de Imágenes para la Evaluación Regulatoria): software utilizado para generar modelos digitales del seno, conocidos como fantasmas, que permiten la evaluación de técnicas de imagen mamográfica.

Fantasma Mamario: representación digital del seno utilizada en la simulación y estudio de imágenes mamográficas para evaluar y mejorar las técnicas de imagen.

Espacio Aéreo: espacio vacío entre las paletas de compresión y el tejido mamario que puede causar variaciones no deseadas en la intensidad de la imagen mamográfica.

Corrección de Intensidad: método utilizado para ajustar las variaciones de intensidad en una imagen digital, asegurando una representación más precisa del tejido.

Resumen

Título: Diseño e implementación de un algoritmo para la estimación y corrección de la deformación del seno en mamografías digitales mediante procesamiento de imágenes. *

Autor: Luis Alejandro Pedreros Martínez y Paula Andrea Mora Escudero **

Palabras Clave: Cáncer de seno, mamografía digital, corrección de la deformación, procesamiento de imágenes médicas, compresión del tejido mamario, seno.

Descripción: La mamografía es la herramienta principal para la detección temprana del cáncer de seno, una de las principales causas de mortalidad en mujeres a nivel mundial. Durante la adquisición de imágenes mamográficas, el seno es comprimido por dos paletas, y dicha compresión introduce deformaciones que resultan en variaciones de intensidad no deseadas, afectando la calidad de la imagen y complicando la detección precisa de lesiones. Este trabajo presenta un modelo automático para la detección y corrección de estas variaciones de intensidad, utilizando un enfoque basado en el análisis del histograma de la imagen. La metodología incluye un sistema de detección automática del *air gap* (espacio de aire entre el seno y las paletas de compresión), seguido de un modelo exponencialmente decreciente diseñado para corregir las variaciones de intensidad en las mamografías digitales. La efectividad de este enfoque se evaluó mediante experimentos con 30 *phantoms* mamarios virtuales generados con el software VICTRE, que simulan diversas condiciones clínicas y distribuciones de tejido mamario. Los resultados indican una mejora significativa en la calidad de las imágenes procesadas, con un incremento en el índice de similitud estructural (SSIM) de 0,946 a 0,996 y una reducción en el error cuadrático medio (RMSE) de 2,51 a 2,03. Estos hallazgos subrayan la importancia de aplicar técnicas de corrección de deformaciones para mejorar la calidad de las imágenes mamográficas.

* Trabajo de Grado

** Facultad de Ingenierías Fisicomecánicas. Escuela de Ingeniería Eléctrica, Electrónica y de Telecomunicaciones.
Director: Said David Pertuz Arroyo. Co-Directora: Mónica Andrea Botero Londoño

Abstract

Title: Automatic Correction of Intensity Variations due to Breast Deformation in Digital Mammograms. *

Author: Luis Alejandro Pedreros Martínez and Paula Andrea Mora Escudero **

Key Words: Breast cancer, digital mammography, deformation correction, medical image processing, breast tissue compression, breast.

Description: Mammography is the main tool for the early detection of breast cancer, one of the leading causes of mortality in women worldwide. During mammographic image acquisition, the breast is compressed by two paddles, and such compression introduces deformations that result in unwanted intensity variations, affecting image quality and complicating accurate lesion detection. This paper presents an automatic model for the detection and correction of these intensity variations, using an approach based on image histogram analysis. The methodology includes an automatic air gap detection system, followed by an exponentially decreasing model designed to correct intensity variations in digital mammograms. The effectiveness of this approach was evaluated by experiments with 30 virtual breast phantoms generated with VICTRE software, simulating various clinical conditions and breast tissue distributions. The results indicate a significant improvement in the quality of the processed images, with an increase in the structural similarity index (SSIM) from 0.946 to 0.996 and a reduction in the root mean square error (RMSE) from 2.51 to 2.03. These findings underscore the importance of applying deformation correction techniques to improve the quality of mammographic images.

* Undergraduate Work

** Faculty of Physicomechanical Engineering. School of Electrical, Electronic and Telecommunications Engineering. Advisor: Said David Pertuz Arroyo. Co-Advisor: Mónica Andrea Botero Londoño

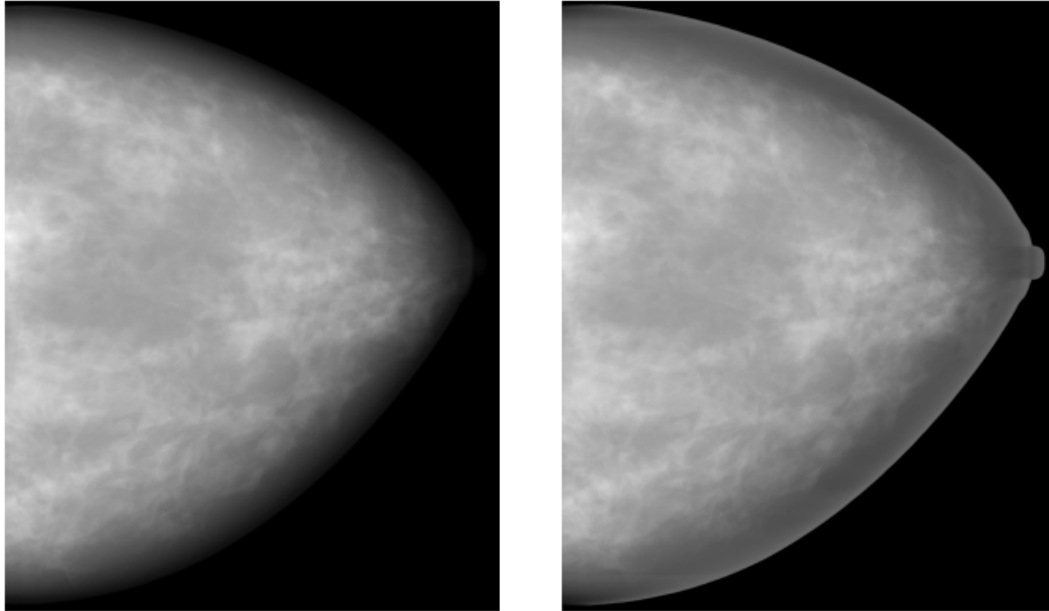
Introducción

El cáncer de mama sigue siendo una de las principales causas de mortalidad en mujeres a nivel mundial, constituyendo un desafío significativo para la salud pública. Según la Organización Mundial de la Salud (OMS), alrededor de 7.8 millones de mujeres han sido diagnosticadas con esta enfermedad en los últimos años, lo que subraya la urgente necesidad de estrategias efectivas para la detección temprana WHO, 2020. La mamografía, debido a su accesibilidad, bajo costo y capacidad para detectar lesiones en etapas tempranas, se ha consolidado como la herramienta principal para la detección del cáncer de mama Koriech, 1996. Sin embargo, este método no está exento de desafíos técnicos significativos.

Durante la adquisición de imágenes mamográficas, el seno se comprime entre dos paletas para aplanar el tejido mamario y mejorar la resolución de la imagen. Aunque esta compresión es necesaria para obtener imágenes claras, introduce deformaciones en el tejido mamario que resultan en variaciones no deseadas en la intensidad de la imagen. Estas variaciones afectan la calidad de la imagen y complican la detección precisa de lesiones, particularmente en áreas donde la densidad del tejido mamario es alta Chougrad et al., 2018. Este desafío es especialmente crítico dado que la capacidad de la mamografía para detectar anomalías en etapas tempranas depende en gran medida de la uniformidad y precisión de la imagen obtenida. La inclinación de la paleta de compresión provoca variaciones en el grosor del seno, resultando en inhomogeneidades en la intensidad de la mamografía; corregir estas deformaciones es esencial para mejorar la precisión diagnóstica. Este efecto se ilustra en la Figura 1.

Figura 1

Efecto de la deformación mamaria en las intensidades de imagen



Nota: Izquierda: imagen original. Derecha: imagen después de corregir las variaciones de intensidad espacial.

La evolución de la tecnología mamográfica ha estado marcada por continuos esfuerzos para mejorar la calidad de las imágenes y superar estos desafíos técnicos. Desde los primeros días de la radiografía hasta la introducción de la mamografía digital, se han implementado varias innovaciones para mejorar la claridad y precisión de las imágenes mamográficas. La transición de la mamografía analógica a la digital representó un avance significativo, permitiendo una mejor resolución y la posibilidad de realizar un post-procesamiento de las imágenes. Sin embargo, la compresión mamaria sigue siendo un procedimiento estándar que, aunque mejora la nitidez de la imagen, introduce problemas de distorsión que deben abordarse Mitra and Bankman, 2010b;

Siddiqui, 2019.

En la literatura científica, se han documentado varios intentos para corregir estas deformaciones inducidas por la compresión. Kallenberg y Karssemeijer, por ejemplo, han desarrollado métodos para corregir la inclinación en mamografías digitales utilizando técnicas de detección de bordes y optimización Kallenberg and Karssemeijer, 2010, 2012. Otros estudios han propuesto modelos de parámetros globales para corregir el grosor del tejido mamario, mejorando la uniformidad de la intensidad en las imágenes Snoeren and Karssemeijer, 2004. Sin embargo, muchas de estas soluciones dependen de datos específicos del proceso de adquisición, como la inclinación de las paletas de compresión o la fuerza aplicada, que no siempre están disponibles para el procesamiento posterior de las imágenes.

En respuesta a estas limitaciones, el presente trabajo se centra en el desarrollo de un modelo automático que no depende de datos adicionales del proceso de adquisición, sino que utiliza únicamente la información presente en la imagen para detectar y corregir las variaciones de intensidad inducidas por la deformación del seno. Este enfoque se basa en un análisis exhaustivo del histograma de la imagen para identificar el *Air gap*, seguido de un modelo de disminución exponencial que ajusta las variaciones de intensidad. Para validar la efectividad de este modelo, se realizaron experimentos con 30 *phantoms* virtuales de mama generados con el software VICTRE, lo que permitió una evaluación precisa de la calidad de la estimación de la deformación y la efectividad de la corrección de las variaciones de intensidad DIDSR, 2024a.

Los *phantoms* virtuales son representaciones digitales precisas del seno que permiten la simulación de diferentes condiciones clínicas y tipos de tejido mamario. Estos modelos son esenciales para la investigación en mamografía, ya que proporcionan un entorno controlado para evaluar nuevos algoritmos y técnicas. En este estudio, se utilizaron *phantoms* para generar mamografías con deformaciones conocidas, lo que permitió evaluar objetivamente la efectividad del algoritmo propuesto.

1. Antecedentes

En esta sección se revisarán los fundamentos teóricos y conceptuales relacionados con la corrección automática de las variaciones de intensidad causadas por la deformación mamaria en la mamografía digital. Se discutirán los conceptos clave de la mamografía digital, incluyendo su importancia técnica y clínica, así como los avances tecnológicos que han mejorado significativamente la calidad de la imagen y la precisión diagnóstica. Además, se abordarán los desafíos relacionados con las inhomogeneidades de intensidad, es decir, las variaciones de imagen debidas a la compresión no uniforme del seno, y cómo estos desafíos impactan la interpretación de las imágenes. También se revisarán estudios previos que han propuesto diversas metodologías para corregir estas deformaciones, analizando su efectividad, limitaciones y aplicaciones prácticas. Finalmente, se justificará el enfoque propuesto en esta investigación, explicando cómo se basa en técnicas existentes para ofrecer mejoras en la precisión y eficiencia en la corrección de imágenes mamográficas, estableciendo una base sólida para la justificación del enfoque propuesto.

1.1. Conceptos Clave

Para entender adecuadamente el enfoque de esta investigación, es esencial definir varios conceptos clave utilizados en el análisis y procesamiento de mamografías.

1.1.1. Mamografía Digital

La mamografía digital es una técnica de imagen que utiliza rayos X de baja dosis para obtener imágenes detalladas del interior del seno. Estas imágenes se digitalizan para su análisis y almacenamiento, lo que permite una manipulación mejorada y la aplicación de algoritmos avanzados de procesamiento de imágenes que mejoran la detección de anomalías. En comparación con la mamografía analógica, la mamografía digital ofrece una mayor resolución de contraste, la capacidad de ajustar el brillo y el contraste de la imagen, y la posibilidad de aplicar técnicas de procesamiento para resaltar detalles específicos.

1.1.2. Compresión Mamaria

Durante la mamografía, el seno se comprime entre dos placas para aplanar el tejido y obtener imágenes más claras y detalladas. Esta compresión reduce el grosor del tejido a través del cual deben pasar los rayos X, minimiza la dosis de radiación requerida y mejora la calidad de la imagen al reducir la dispersión de los rayos X. Sin embargo, la inclinación de la paleta de compresión puede causar variaciones en el grosor del seno, resultando en inhomogeneidades de intensidad en la mamografía; corregir estas deformaciones es esencial para mejorar la precisión diagnóstica.

1.1.3. *Inhomogeneidades de Intensidad*

Las inhomogeneidades de intensidad se refieren a variaciones en la intensidad de la imagen debido a la distribución desigual del tejido mamario durante la compresión. Estas variaciones pueden ocultar o simular lesiones, lo que representa un gran desafío en la interpretación de mamografías y puede afectar la precisión diagnóstica.

1.1.4. *Histograma de una Imagen*

Un histograma es una representación gráfica de la distribución tonal en una imagen digital. En el contexto de la mamografía, el histograma proporciona información sobre la distribución de la intensidad de la imagen, crucial para entender la composición subyacente del tejido. Al analizar el histograma, se pueden identificar regiones de alta densidad (tejido fibroglandular) y baja densidad (tejido adiposo). Las variaciones en el histograma pueden indicar inhomogeneidades de intensidad causadas por una compresión desigual del seno durante la toma de imágenes. El histograma es una herramienta vital para las técnicas de mejora y normalización de imágenes.

1.1.5. *Índice de Similitud Estructural (SSIM)*

El índice de similitud estructural (SSIM) es una métrica utilizada para evaluar la similitud entre dos imágenes, considerando factores como el brillo, el contraste y la estructura. El SSIM es importante para evaluar la calidad de las imágenes corregidas y compararlas con las imágenes originales, ya que proporciona una medida cuantitativa de la similitud visual entre imágenes.

1.1.6. Error Cuadrático Medio (RMSE)

El error cuadrático medio (RMSE) mide la diferencia promedio entre los valores predichos por un modelo y los valores observados. Es crucial para cuantificar la efectividad de los algoritmos de corrección de intensidad, donde un RMSE más bajo indica una mayor precisión en la corrección de deformaciones.

1.2. Revisión de Estudios Previos

Esta sección revisa la literatura existente sobre los métodos y enfoques utilizados para la corrección de deformidades en mamografías.

1.2.1. Modelos de Parámetros Globales

Los modelos de parámetros globales ajustan las imágenes mamográficas basándose en parámetros conocidos del proceso de adquisición, como la inclinación de la paleta de compresión y la distribución del grosor del tejido mamario. Snoeren y Karssemeijer (2004) propusieron un modelo de parámetros globales para corregir el grosor de mamografías comprimidas, mostrando mejoras significativas en la calidad de la imagen. Sin embargo, estos enfoques tienen limitaciones, ya que dependen de datos específicos del equipo de adquisición y pueden no ser aplicables en todos los casos.

1.2.2. Algoritmos de Detección de Bordos

Los algoritmos de detección de bordes identifican los límites del tejido mamario en las imágenes mamográficas y corrigen las deformaciones causadas por la compresión desigual. Kallenberg²⁰¹⁰,kallenberg²⁰¹²compression desarrollaron métodos para corregir la inclinación en mamografías digitales utilizando técnicas de detección de bordes y optimización. Estos métodos han demostrado ser efectivos, pero dependen de información específica del proceso de adquisición, como la inclinación de la paleta de compresión o la fuerza de compresión, que generalmente está disponible en los sistemas mamográficos, pero no es adecuada para el procesamiento posterior de mamografías digitales.

1.3. Métodos Actuales para la Corrección de Deformaciones

Los métodos actuales para la corrección de deformaciones en mamografías se dividen en dos categorías principales: aquellos basados en modelos de parámetros globales y algoritmos de detección de bordes.

1.3.1. Métodos Basados en Modelos de Parámetros Globales

Los métodos basados en modelos de parámetros globales ajustan la imagen mamográfica basándose en parámetros conocidos del proceso de adquisición, mejorando la uniformidad de la imagen y reduciendo las variaciones de intensidad. Sin embargo, su dependencia de datos específicos del equipo de adquisición puede limitar su aplicabilidad en diferentes entornos clínicos. Estos

métodos son efectivos para mejorar la calidad de la imagen, pero pueden no ser adecuados para todas las situaciones clínicas debido a la variabilidad en el equipo de adquisición y las condiciones de imagen.

1.3.2. Algoritmos de Detección de Bordes

Los algoritmos de detección de bordes utilizan técnicas de optimización para identificar y corregir deformaciones en imágenes mamográficas. Estos métodos han demostrado ser efectivos para mejorar la visibilidad de las estructuras internas del seno y la precisión diagnóstica, aunque también pueden requerir ajustes específicos para diferentes tipos de tejido mamario. Los algoritmos de detección de bordes son especialmente útiles para corregir deformaciones localizadas y mejorar la precisión de la imagen en áreas específicas.

1.4. Justificación y Ventajas del Enfoque Propuesto

La corrección automática de variaciones de intensidad en mamografías digitales es esencial para mejorar la precisión diagnóstica del cáncer de mama. Los métodos actuales, aunque útiles, tienen limitaciones significativas, como la dependencia de datos específicos del equipo de adquisición y la necesidad de ajustes para diferentes tipos de tejido mamario. El enfoque propuesto corrige las variaciones de intensidad sin requerir información específica del proceso de adquisición, lo que lo hace ampliamente aplicable en diferentes entornos clínicos.

Este capítulo ha revisado la evolución de la mamografía y los avances tecnológicos que

han mejorado la calidad de las imágenes y la precisión diagnóstica. Se han definido conceptos clave como la mamografía digital, la compresión mamaria y las inhomogeneidades de intensidad. También se han revisado los métodos de corrección de deformaciones, destacando los modelos de parámetros globales y los algoritmos de detección de bordes, con sus respectivas ventajas y limitaciones. El enfoque propuesto en esta investigación, que corrige automáticamente las variaciones de intensidad, promete mejorar la precisión diagnóstica y reducir la carga de trabajo de los radiólogos. Esta revisión proporciona una base sólida para el desarrollo de métodos avanzados de corrección de imágenes mamográficas.

2. Materiales y Métodos

Este capítulo presenta los materiales y métodos utilizados para abordar la corrección automática de variaciones de intensidad en mamografías digitales, derivadas de la deformación del seno durante la compresión. Se describe la creación de una base de datos de *phantoms* virtuales de mama, diseñados para replicar diversas condiciones clínicas y tipos de tejido mamario. Además, se presenta el desarrollo de un algoritmo que estima la deformación y corrige las variaciones de intensidad en las imágenes. Este proceso incluye el preprocesamiento de imágenes, la detección de áreas afectadas por la compresión y la aplicación de un modelo de compensación de intensidad. Finalmente, se detallan las métricas utilizadas para evaluar la efectividad del método, centrándose en mejorar la precisión diagnóstica de las mamografías.

2.1. Base de Datos

Para este estudio, se utilizó el software *Virtual Imaging Clinical Trial for Regulatory Evaluation* (VICTRE), desarrollado por la División de Imágenes, Diagnóstico y Fiabilidad de Software (DIDSR), para simular ensayos clínicos de imágenes. VICTRE permite la creación de *Virtual Breast Phantoms* (VBPs), que son representaciones digitales detalladas del seno con características anatómicas predefinidas. Este software es esencial para generar bases de datos controladas que permiten la evaluación y validación de técnicas de procesamiento de imágenes, como la corrección de variaciones de intensidad en mamografías DIDSR, 2024a.

La base de datos generada consiste en 30 *phantoms* virtuales de mama, cada uno acom-

pañado de su correspondiente mamografía digital. Estos *phantoms* fueron diseñados para simular diversas condiciones clínicas y tipos de tejido mamario, asegurando que las imágenes representen fielmente las características anatómicas y patológicas relevantes para el análisis DIDSr, 2024b.

2.1.1. Generación de *Phantoms* y Mamografías

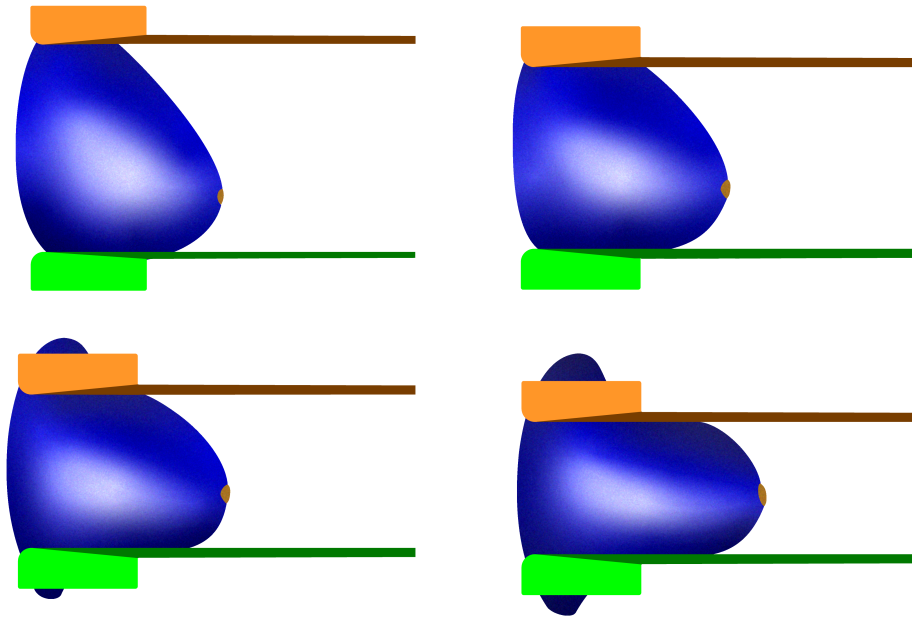
La generación de VBPs es un proceso crítico que comienza con un modelo analítico que crea *phantoms* de mama basados en variaciones clave en las características del paciente, como la forma del seno, la glandularidad y la densidad. Cada *phantom* está asociado con una mamografía digitalmente generada que refleja la deformación del tejido mamario bajo compresión, simulando con precisión las condiciones de adquisición de mamografías reales DIDSr, 2024b.

Para replicar este proceso, se utilizó el software *VICTRE breastCompress*, que simula la compresión física que ocurre durante una mamografía. Esta simulación incluye la creación de una malla tetraédrica que modela el comportamiento mecánico del tejido mamario bajo presión, permitiendo que el *phantom* refleje con precisión la morfología resultante de la compresión.

Posteriormente, los *phantoms* y sus mamografías fueron recortados utilizando el software *VICTRE breastCrop* para estandarizar el volumen de análisis y optimizar los recursos computacionales. Este paso asegura que solo las áreas relevantes del seno se incluyan en los análisis posteriores DIDSr, 2024c.

Figura 2

El Air gap es el espacio vacío entre las paletas de compresión y el seno, y es la causa de variaciones de intensidad en la mamografía



2.1.2. Categorías de Densidad Mamaria

Las mamografías generadas se clasificaron según el Sistema de Informes y Datos de Imágenes de Mama (BI-RADS) en las siguientes categorías:

- **Dense breasts:** Alta densidad de tejido fibroglandular.
- **Fatty breasts:** Mayor proporción de tejido adiposo.
- **Heterogeneously dense breasts:** Mezcla de tejido denso y graso.
- **Scattered fibroglandular breasts:** Distribución irregular de densidades.

Esta clasificación asegura que el estudio abarque una amplia gama de variaciones en la densidad mamaria, proporcionando un entorno robusto para evaluar la efectividad del algoritmo de corrección de intensidad DIDSr, 2024b.

2.1.3. Calidad de los Datos y Preprocesamiento

Los VBPs se crearon con un enfoque en la calidad y consistencia de la imagen. Se aplicaron técnicas de preprocesamiento para estandarizar las intensidades, normalizar el brillo y el contraste, y reducir el ruido en las imágenes. Estas técnicas son cruciales para preparar adecuadamente las imágenes para la aplicación del algoritmo propuesto DIDSr, 2024a.

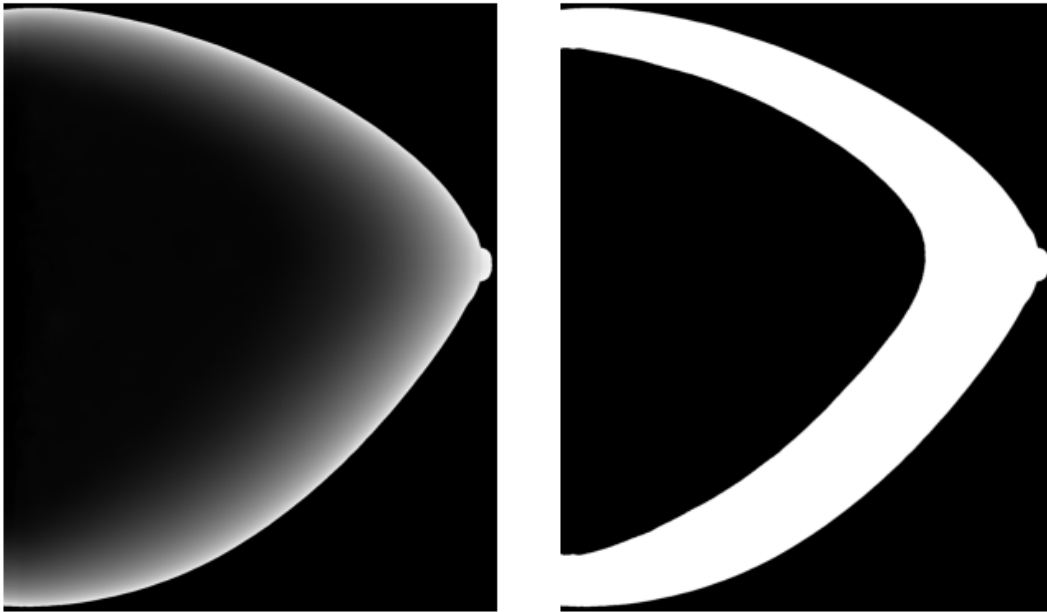
2.2. Generación de Verdades Fundamentales (GT)

Para evaluar la precisión del algoritmo de corrección de intensidad propuesto, es esencial contar con una referencia confiable que permita medir la efectividad de la corrección aplicada. Esta referencia se conoce como *Ground Truth* (GT), y en este contexto, representa la ubicación exacta y precisa del *Air gap* entre las paletas de compresión y el seno deformado durante la adquisición de la mamografía.

El *Air gap* es una región clave en la imagen mamográfica, ya que su correcta identificación es crucial para la aplicación del modelo de corrección de intensidad. El *Air gap* se refiere al vacío que se forma entre las paletas de compresión y la superficie del seno comprimido, lo que resulta en variaciones de intensidad que deben corregirse para obtener una imagen más precisa y útil desde

Figura 3

Generación de Ground Truth



Nota: Izquierda: el mapa de deformación mide el porcentaje de aire en cada posición 2D del plano XY. Derecha: el Air gap se encuentra aplicando un umbral del 20% de aire en el mapa de deformación.

el punto de vista diagnóstico.

Para generar el *Ground Truth* del *Air gap*, seguimos un proceso que involucra el uso de los *phantoms* de mama 3D generados:

2.2.1. Almacenamiento de Rebanadas del Phantom

Las rebanadas del *phantom* se almacenan en una matriz tridimensional de datos enteros. En esta matriz, cada valor representa un tipo específico de tejido o estructura anatómica, lo que nos permite construir un modelo detallado del seno y sus características internas.

2.2.2. Generación del Campo de Deformación 2D

A partir de la matriz 3D, se obtiene un campo de deformación 2D contando los vóxeles que representan tejido mamario a lo largo del eje z. Este campo de deformación 2D refleja la proporción de vóxeles de tejido en comparación con los vóxeles de aire en cada posición 2D de la imagen mamográfica.

2.2.3. Segmentación del Air Gap

Posteriormente, el *Air gap* se segmenta aplicando un umbral específico al campo de deformación 2D. Este umbral, V_{cut} , se determina empíricamente y representa el porcentaje mínimo de vóxeles de aire necesario para clasificar una región como un *Air gap*. En este estudio, se utilizó un valor de corte del 20% para definir el *Air gap* (10% para la zona superior y 10% para la zona inferior).

$$V_{\text{cut}} = \frac{V_{\text{air}}}{H} = 0.2 \quad (1)$$

2.2.4. Creación de la Máscara GT

La segmentación del *Air gap* genera una máscara binaria que indica la ubicación precisa del borde entre el tejido mamario y el aire. Esta máscara sirve como referencia (*Ground Truth*) para evaluar la precisión del algoritmo propuesto de detección del *Air gap*.

El proceso completo de detección del *Air gap* puede expresarse matemáticamente de la

siguiente manera:

$$GT(i, j) = \sum_{k=1}^H \delta(\text{slice}(i, j, k) - \text{tissue_value}) \quad (2)$$

$$\text{Air gap } GT(i, j) = GT(i, j) > \text{Threshold} \quad (3)$$

Donde:

- $\text{slice}(i, j, k)$ es el valor del vóxel en la fila i , columna j y capa k ; H es la altura de la matriz 3D y tissue_value es el valor específico que identifica cada tipo de tejido.
- $\delta(x)$ es la función delta de Kronecker que devuelve 1 si $x = 0$ y 0 en caso contrario.
- $\text{Air gap } GT(i, j)$ es la máscara binaria que indica el borde en la fila i y la columna j .
- V_{cut} es el umbral determinado por la relación entre el porcentaje de vóxeles de aire y la altura del seno.

Esta *Ground Truth* es esencial para validar el algoritmo de corrección de intensidad, ya que proporciona un estándar objetivo contra el cual se comparan los resultados obtenidos por el algoritmo.

2.3. Corrección de Intensidad

El proceso de corrección de intensidad aborda las variaciones en la intensidad de la imagen causadas por la deformación del seno durante la compresión mamográfica. Involucra tres pasos

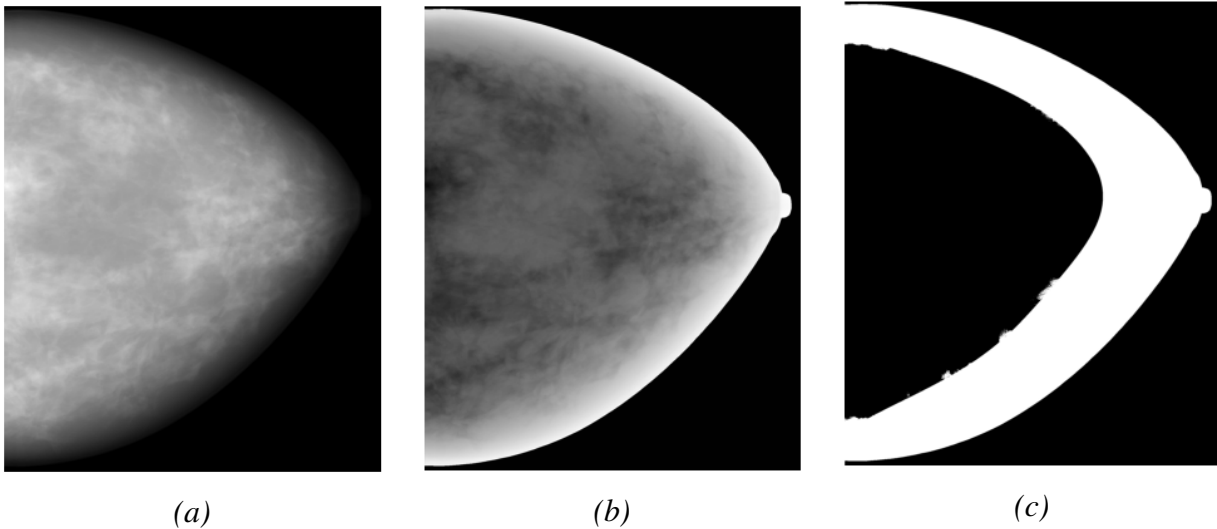
clave: preprocesamiento, detección del *Air gap* y la aplicación de un modelo de corrección de intensidad.

2.3.1. *Preprocesamiento*

En este paso, se estandarizan las imágenes mamográficas, se identifica el fondo y se realiza una inversión de los valores de intensidad de las imágenes para facilitar los pasos subsiguientes (Figura 4).

Figura 4

Pre-procesamiento y detección del Air gap



Nota: (a) imagen original, (b) imagen preprocesada y (c) Air gap detectado.

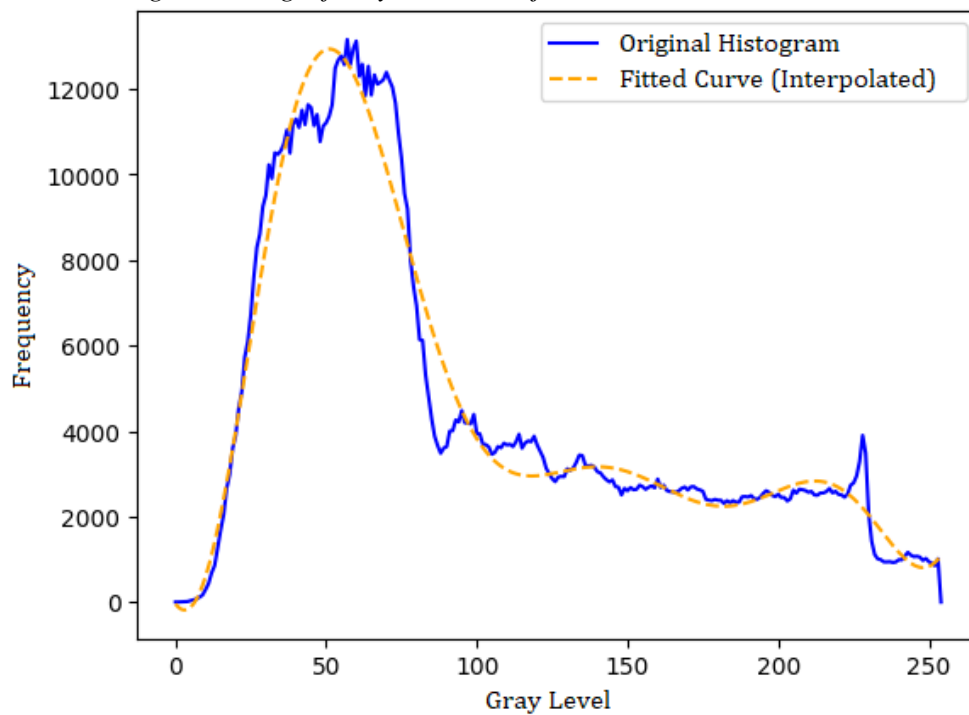
2.3.2. *Detección del Air Gap*

El objetivo de este paso es detectar la región donde la deformación del seno induce cambios en la intensidad de la mamografía (Figura 4). En nuestro enfoque, se obtiene un umbral de

intensidad óptimo a partir del histograma de la imagen mamográfica. La mayoría de las intensidades de la imagen se concentran alrededor de un lóbulo principal que corresponde a los píxeles en la región interna del seno, sin los efectos del *Air gap*. En contraste, los píxeles dentro del *Air gap* corresponden a una zona plana después del lóbulo principal en el histograma (Figura 5). Este cambio abrupto en la distribución de frecuencias se utiliza para encontrar el corte de intensidad óptimo para identificar el *Air gap*.

Figura 5

Histograma de la imagen mamográfica y su curva ajustada



Primero, se calcula el histograma de la imagen preprocesada $I_{pre}(i, j)$ para analizar la distribución de los niveles de gris:

$$H(I) = \text{histogram}(I_{\text{pre}}(i, j)) \quad (4)$$

Luego, se realiza un submuestreo del histograma y se ajusta un polinomio al conjunto de datos submuestreados para reducir el ruido; posteriormente, la curva ajustada se interpola para coincidir con el dominio original del histograma:

$$C(x) = \text{interp} \left(\sum_{k=0}^n a_k x^k \right) \quad (5)$$

Donde $C(x)$ representa la curva interpolada, interp es la función de interpolación, a_k son los coeficientes del polinomio, x es el nivel de intensidad y n es el orden del polinomio.

Finalmente, se encuentra el nivel de gris cuya frecuencia es más cercana al promedio ajustado después de que la curva ha alcanzado su pico:

$$C_{\text{opt}} = \arg \min_x |C(x) - \text{mean}(C(x))| \quad (6)$$

Este procedimiento encuentra el nivel de gris que identifica el borde del seno, ajustándose a las características físicas de cada mamografía.

El C_{opt} obtenido en (6) es sensible al ruido de la distribución de cada mamografía; se obtiene un corte de intensidad óptimo específico para cada categoría de densidad al promediar los cortes de intensidad correspondientes según la categoría de densidad mamaria. Este enfoque tiene

como objetivo reducir el ruido de la estimación del borde. Posteriormente, se encuentra el borde utilizando el corte óptimo C_{opt} y se suaviza ajustando un polinomio que suaviza la curva donde comienza la región no comprimida del seno, como se muestra en la Fig. 4:

$$\text{Air gap}(i, j) = I_{\text{pre}}(i, j) > C_{\text{opt}} \quad (7)$$

2.3.3. Aplicación de la Corrección de Intensidad

El *Air gap* suavizado encontrado en (7) se utiliza como una máscara para la corrección de intensidad en las áreas no comprimidas. Para cada fila i , en la zona no comprimida del seno, se genera un modelo de compensación de intensidad $\mathbf{M}[i, j]$ basado en la propagación de rayos X en una mamografía Snoeren and Karssemeijer, 2004:

$$\mathbf{M}[i, j] = \text{Edge}(i, j) \cdot \exp\left(-\frac{j}{N_i/\alpha}\right) \quad (8)$$

Donde α es un factor de atenuación para la compensación y N_i corresponde al borde no comprimido en la fila i , definido en el rango de columnas j para las cuales se realiza la compensación. Finalmente, se aplica la corrección de intensidad a la imagen original I , normalizada como F :

$$H(i, j) = F(i, j) + g \cdot (1 - F(i, j)) \cdot \mathbf{M}(i, j) \quad (9)$$

Aquí, g es un factor de ganancia ajustable y $H(i, j)$ representa la imagen corregida. Esta

metodología permite identificar y suavizar el borde no comprimido del seno y ajustar la intensidad según la estructura del borde suavizado, mejorando así la visibilidad del tejido mamario y la calidad de las imágenes mamográficas.

2.4. Evaluación de la Corrección de Intensidad

Evaluar la corrección de intensidad es crucial para asegurar que las imágenes corregidas sean de alta calidad y clínicamente útiles. Se emplean varias métricas para evaluar cuantitativamente la efectividad del algoritmo.

Error Cuadrático Medio (RMSE) Mide la diferencia cuadrática media entre la imagen corregida y la imagen de referencia. Un RMSE más bajo indica una mayor coincidencia con la referencia, reflejando una corrección de intensidad más precisa.

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (I_{\text{corr}}(i) - I_{\text{GT}}(i))^2} \quad (10)$$

Índice de Similitud Estructural (SSIM) Evalúa la calidad percibida de las imágenes corregidas comparando la información estructural, la luminancia y el contraste entre la imagen corregida y la imagen de referencia. Los valores de SSIM varían de -1 a 1, siendo los valores más cercanos a 1 los que indican mejor similitud y calidad.

$$\text{SSIM} = \frac{(2\mu_{I_{\text{corr}}} \mu_{I_{\text{GT}}} + C_1)(2\sigma_{I_{\text{corr}}, I_{\text{GT}}} + C_2)}{(\mu_{I_{\text{corr}}}^2 + \mu_{I_{\text{GT}}}^2 + C_1)(\sigma_{I_{\text{corr}}}^2 + \sigma_{I_{\text{GT}}}^2 + C_2)} \quad (11)$$

La evaluación incluye tanto la inspección visual como la comparación cuantitativa entre las imágenes corregidas y los datos de referencia. La inspección visual permite identificar cualquier artefacto o distorsión residual que pueda no ser capturado por las métricas cuantitativas. Mientras tanto, las métricas cuantitativas mencionadas anteriormente, calculadas para imágenes corregidas y no corregidas, resaltan las mejoras en la calidad de la imagen, demostrando la efectividad del algoritmo para reducir las variaciones de intensidad y mejorar la visibilidad de características críticas.

Finalmente, el uso de datos de referencia, que incluyen modelos 3D precisos del seno, permite una validación exhaustiva del proceso de corrección. Al comparar las imágenes corregidas con estos modelos, aseguramos que el proceso de corrección refleje con precisión las características anatómicas reales y reduzca el impacto de los artefactos de compresión.

3. Experimentos y Resultados

En esta sección, presentamos los experimentos realizados y los resultados obtenidos en la evaluación del algoritmo propuesto para la corrección automática de variaciones de intensidad en mamografías digitales. Los experimentos se llevaron a cabo en dos partes principales: Primero, medimos el rendimiento en la detección automática del *Air gap*. Segundo, medimos la calidad de la imagen con la corrección de intensidad.

3.1. Detección del Air Gap

La detección del *Air gap* es un paso crucial para identificar los píxeles que requieren corrección de intensidad. Este *gap* se refiere al área vacía entre las paletas de compresión y el seno deformado durante la adquisición de la mamografía. Para evaluar la precisión de la detección del *Air gap*, se comparó la *Ground Truth* con el *Air gap* detectado utilizando el método propuesto. Específicamente, se midió la distancia promedio en píxeles entre el borde del *Air gap* de la *Ground Truth* y el borde detectado por nuestro modelo. Los resultados se resumen en la Tabla 1. Para calcular SSIM y RMSE, utilizamos como referencia la imagen con corrección de intensidad utilizando el mapa de deformación de la *Ground Truth* (ver Fig. 6).

Los resultados mostraron que el algoritmo propuesto es capaz de detectar el *Air gap* con una precisión razonable. Los senos grasos mostraron el menor error promedio, mientras que los senos densos y heterogéneos mostraron un mayor error, lo que puede deberse a la mayor complejidad en

Tabla 1. *Distancia promedio en píxeles en el Air gap detectado*

<i>Tipo</i>	<i>Distancia Promedio %</i>	<i>STD</i>
<i>Dense</i>	9.615	13.598
<i>Fatty</i>	3.649	2.957
<i>Hetero</i>	9.224	5.928
<i>Scatter</i>	8.166	6.788
<i>Promedio</i>	8.234	6.487

la detección de bordes en estos tipos de tejido.

3.2. Corrección de Intensidad

Para evaluar la corrección de intensidades en la imagen mamográfica, se aplicó un modelo de corrección generado a partir de la *Ground Truth*. Luego, esta corrección se comparó con la corrección obtenida por el método propuesto. La calidad de la corrección se evaluó utilizando dos métricas principales: el índice de similitud estructural (SSIM) y el error cuadrático medio (RMSE), ver Tablas 2 y 3.

Tabla 2. *Calidad de la corrección de intensidad usando SSIM*

<i>Tipo</i>	<i>Sin Corrección</i>	<i>Con Corrección</i>
<i>Dense</i>	0.934 (0.0003)	0.996 (0.0001)
<i>Fatty</i>	0.946 (0.0030)	0.997 (0.0011)
<i>Hetero</i>	0.943 (0.0021)	0.996 (0.0005)
<i>Scatter</i>	0.951 (0.0030)	0.996 (0.0009)
<i>Promedio</i>	0.946 (0.0054)	0.996 (0.0008)

Los resultados indican una mejora significativa en la precisión de las imágenes después de

Tabla 3. *Calidad de la corrección de intensidad usando RMSE*

<i>Tipo</i>	<i>Sin Corrección</i>	<i>Con Corrección</i>
<i>Dense</i>	2.593 (0.0413)	1.741 (0.3460)
<i>Fatty</i>	2.560 (0.1815)	1.980 (0.4222)
<i>Hetero</i>	2.534 (0.0141)	2.051 (0.2246)
<i>Scatter</i>	2.479 (0.0286)	2.070 (0.1889)
<i>Promedio</i>	2.518 (0.0255)	2.031 (0.2392)

la corrección. Sin corrección, los RMSE varían entre 2.4791 y 2.5932, mientras que con la corrección utilizando el modelo propuesto, los RMSE se reducen significativamente a valores entre 1.7407 y 2.0697. El SSIM también mejora de manera significativa, lo que sugiere que el algoritmo no solo mejora la calidad visual de las mamografías, sino que también podría mejorar la precisión diagnóstica al facilitar una interpretación más clara de las imágenes.

3.3. Análisis de Resultados

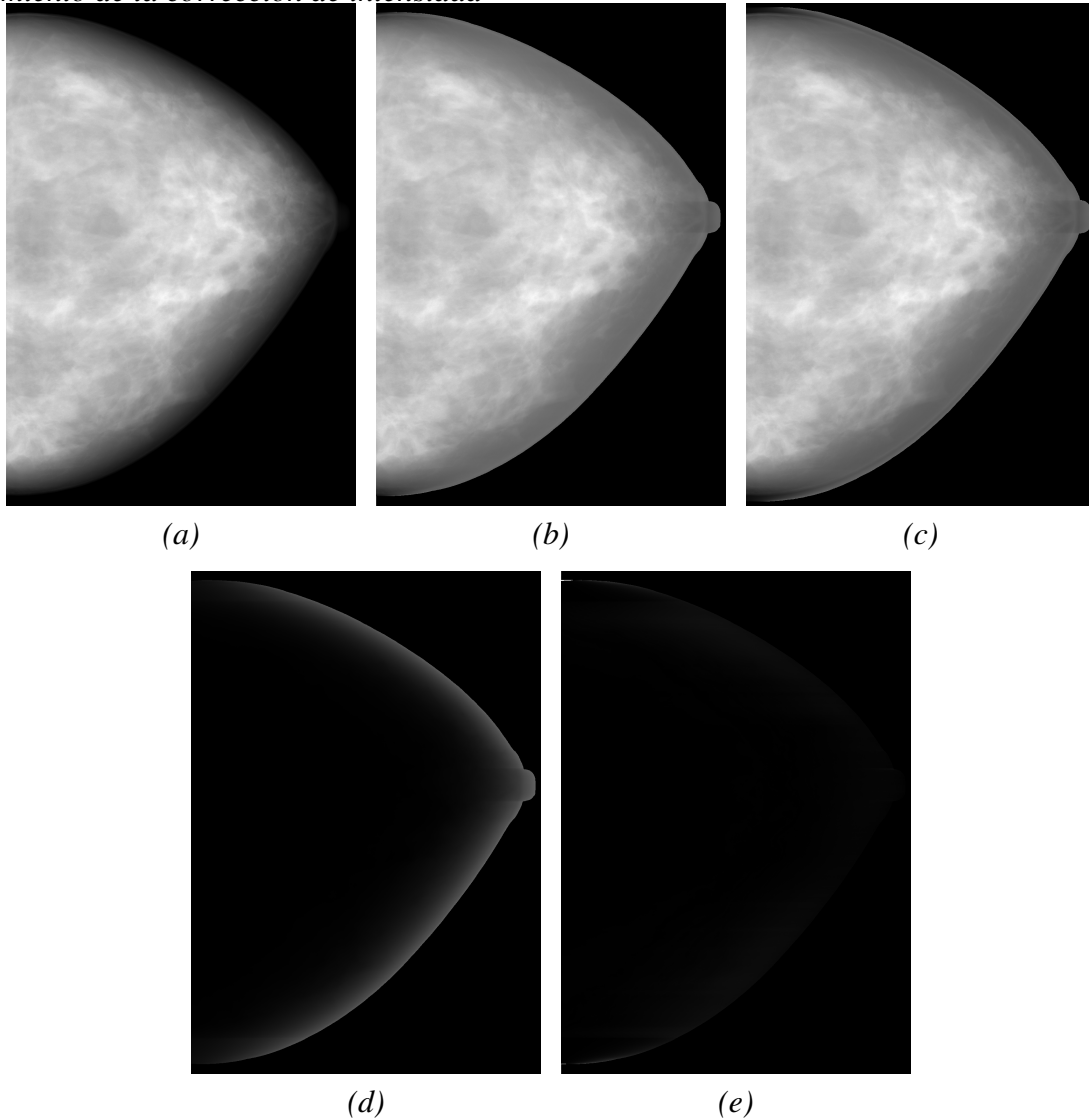
Los resultados obtenidos demuestran una mejora significativa en la calidad de las imágenes mamográficas tras la aplicación del algoritmo propuesto. El índice de similitud estructural (SSIM) aumentó de un promedio de 0.9461 en las imágenes sin corregir a 0.9962 en las imágenes corregidas. Este aumento sugiere que las imágenes corregidas son mucho más similares a las imágenes ideales, lo que indica una mejor preservación de las estructuras anatómicas esenciales.

Además, el error cuadrático medio (RMSE) disminuyó de un promedio de 2.5176 en las imágenes sin corregir a 2.0314 en las imágenes corregidas. Esta reducción en el RMSE refleja una disminución en las diferencias de intensidad entre las imágenes corregidas e ideales, lo que sugiere que el algoritmo no solo mejora la apariencia visual, sino que también reduce los errores

de intensidad que podrían comprometer la precisión diagnóstica.

Figura 6

Rendimiento de la corrección de intensidad



Nota: (a) Imagen original, (b) Corrección de intensidad usando Ground Truth, (c) Corrección de intensidad usando el método propuesto, (d) diferencia entre (a) y (b), y (e) diferencia entre (b) y (c).

En la detección del *Air gap*, los senos grasos mostraron el menor error promedio, lo que indica que el algoritmo es particularmente efectivo en casos con menor complejidad en la detección

de bordes. Sin embargo, en senos densos y heterogéneos, donde la detección de bordes es más desafiante, el error promedio fue mayor. Esto sugiere que, aunque el algoritmo es robusto, puede requerir ajustes adicionales para optimizar su rendimiento en tejidos más complejos.

Para ilustrar estos resultados, se presenta una imagen comparativa que muestra la corrección de intensidad (ver Fig. 6).

La imagen comparativa refuerza visualmente la efectividad del algoritmo propuesto. Las notables diferencias entre las imágenes corregidas y no corregidas, así como la similitud con las imágenes ideales, subrayan la capacidad del algoritmo para mejorar la calidad de las mamografías.

4. Conclusiones

En este estudio, hemos desarrollado y validado un algoritmo para corregir las variaciones de intensidad en mamografías digitales causadas por la deformación del seno durante la compresión. El modelo propuesto se basa en la detección automatizada del *Air gap* directamente desde mamografías 2D y la posterior aplicación de un modelo de corrección exponencial. Es importante destacar que este algoritmo opera utilizando únicamente la información contenida en la imagen, sin requerir datos adicionales sobre el proceso de adquisición, lo que aumenta su versatilidad y aplicabilidad en diferentes entornos clínicos.

Los resultados experimentales mostraron que el algoritmo detecta eficazmente el *Air gap* y corrige las variaciones de intensidad, lo cual se refleja en un aumento significativo del índice de similitud estructural (SSIM) y una reducción del error cuadrático medio (RMSE). En particular, los senos grasos mostraron un menor error en la detección, mientras que en los senos densos y heterogéneos, la mayor complejidad en la detección de bordes resultó en un error ligeramente mayor. Estas mejoras en las métricas de calidad de la imagen validan el potencial del algoritmo para mejorar la calidad de las imágenes mamográficas.

En resumen, el enfoque propuesto ha demostrado ser efectivo en la consecución de los objetivos planteados, mejorando significativamente la calidad de las mamografías digitales al corregir las variaciones de intensidad causadas por la deformación del seno en las mamografías digi-

tales. Los resultados experimentales validan la efectividad del algoritmo, logrando los objetivos de mejora en métricas de calidad de imagen como el índice de similitud estructural (SSIM) y el error cuadrático medio (RMSE).

5. Recomendaciones

A medida que avanzamos en la validación y desarrollo del algoritmo propuesto, se identifican varias áreas clave para futuros trabajos que podrían extender y fortalecer los resultados obtenidos en este estudio.

En primer lugar, se recomienda realizar estudios subjetivos en colaboración con especialistas en radiología e imagen médica. Estos estudios deben involucrar la evaluación del algoritmo utilizando mamografías clínicas reales, en lugar de imágenes sintéticas. La participación de radiólogos permitirá una evaluación cualitativa de la efectividad del algoritmo en condiciones prácticas, proporcionando una perspectiva clínica esencial para complementar los resultados cuantitativos obtenidos hasta ahora. La aplicación del algoritmo en un entorno clínico real es crucial para determinar su impacto en la práctica radiológica diaria y su potencial para mejorar la calidad diagnóstica.

Además, es fundamental validar el algoritmo con una variedad de sistemas mamográficos. Las diferencias en la tecnología de los fabricantes, así como en el preprocesamiento de las imágenes, pueden influir significativamente en los resultados del algoritmo. La evaluación del método en diferentes sistemas permitirá determinar su robustez y adaptabilidad a diferentes entornos tecnológicos. Este paso es esencial para asegurar que el algoritmo funcione de manera consistente y efectiva, independientemente de las variaciones en el equipo y los procedimientos utilizados en diferentes instalaciones clínicas.

Bibliography

- Bankman, I. N. (2000). *Handbook of medical imaging: Processing and analysis*. Academic Press.
- Broeders, M. J. M., et al. (2014). Comparison of a flexible versus a rigid breast compression paddle: Pain experience, projected breast area, radiation dose and technical image quality.
- Chougrad, H., Hamid, Z., & Alheyane, O. (2018). Deep convolutional neural networks for breast cancer screening. *Computer Methods and Programs in Biomedicine*, 157. <https://doi.org/10.1016/j.cmpb.2018.01.011>
- DIDSR. (2024a). Victre breastcompress [Accessed: August 22, 2024]. https://github.com/DIDSR/VICTRE_PIPELINE
- DIDSR. (2024b). Victre pipeline [Accessed: August 22, 2024]. https://github.com/DIDSR/VICTRE_PIPELINE
- DIDSR. (2024c). Victre: Virtual imaging clinical trial for regulatory evaluation [Accessed: August 22, 2024]. <https://github.com/DIDSR/VICTRE>
- Gonzalez, R. C., Woods, R. E., & Eddins, S. L. (2008). *Digital image processing*. Pearson Education India.
- Ikeda, D., & Miyake, K. K. (2017). *Breast imaging: The requisites*. Elsevier.
- Kallenberg, M., & Karssemeijer, N. (2010). Comparison of tilt correction methods in full field digital mammograms. *6136*, 191–196. https://doi.org/10.1007/978-3-642-13666-5_26
- Kallenberg, M., & Karssemeijer, N. (2012). Compression paddle tilt correction in full-field digital mammograms. *Radiology Today*, 15(2), 56–69.

- Karssemeijer, N., & Thijssen, M. (2018). *Digital mammography: A holistic approach*. Springer.
- Kassimi, S., Moussa, H., & Sabiki, H. (2024). Enhancing image denoising: A novel non-local anisotropic diffusion framework based on caputo derivatives and gaussian convolution for the perona–malik model. *Signal Processing*, 222, 109521. <https://doi.org/10.1016/j.sigpro.2024.109521>
- Koriech, O. M. (1996). Breast cancer and early detection. *Journal Family Community Medicine*, 3(1), 7–9.
- Lei, Y.-M., Yin, M., Yu, M.-H., Yu, J., Zeng, S.-E., Lv, W.-Z., li, J., Ye, H.-R., Cui, X.-W., & Dietrich, C. (2021). Artificial intelligence in medical imaging of the breast. *Frontiers in Oncology*, 11, 600557. <https://doi.org/10.3389/fonc.2021.600557>
- Mitra, S. K., & Bankman, I. N. (2010a). *Medical image processing: Techniques and applications*. CRC Press.
- Mitra, S. K., & Bankman, I. N. (2010b). *Medical image processing: Techniques and applications*. CRC Press.
- Otro Autor, E., et al. (2019). Impact of paddle tilt correction in clinical full-field digital mammography. *European Radiology*, 29(3), 1442–1449.
- Siddiqui, S. (2019). Breast cancer detection and diagnosis: A review of the literature.
- Snoeren, P., & Karssemeijer, N. (2004). Thickness correction of mammographic images by means of a global parameter model of the compressed breast. *Medical Imaging, IEEE Transactions on*, 23, 799–806. <https://doi.org/10.1109/TMI.2004.827477>

Timp, S., et al. (2016). Breast compression in mammography in population-based breast cancer screening. *Acta Radiológica*, 57(4), 459–465.

Tschumperlé, D., & Deriche, R. (2017). *Image processing and analysis with graphs: Theory and practice*. CRC Press.

WHO. (2020). Cáncer de seno.

Zheng, Y., et al. (2011). Digital mammography: Effects of image processing on observer performance. *European Radiology*, 21(9), 2018–2025.

Apéndice

Apéndice 1. Base de Datos

Para la validación del algoritmo propuesto, se utilizó un conjunto de datos sintetizado mediante el software *Virtual Imaging Clinical Trial for Regulatory Evaluation* (VICTRE), desarrollado por la *Division of Imaging, Diagnostics, and Software Reliability* (DIDSR). Este conjunto de datos fue esencial para asegurar un entorno controlado en el que evaluar la precisión y eficacia del algoritmo en la corrección de variaciones de intensidad en mamografías digitales DIDSR, 2024a.

Composición de la Base de Datos

La base de datos está compuesta por 30 *phantoms* virtuales de mama. Cada *phantom* incluye un modelo 3D altamente detallado del seno comprimido, así como su correspondiente mamografía digital 2D. Estas imágenes fueron diseñadas para imitar de cerca las condiciones clínicas y anatómicas del seno durante un examen mamográfico DIDSR, 2024b.

Categorización de la Densidad Mamaria

Los *phantoms* fueron categorizados según el *Breast Imaging Reporting and Data System* (BI-RADS) en cuatro tipos principales:

- *Dense*: Senos con alta densidad de tejido fibroglandular.

- *Fatty*: Senos con una mayor proporción de tejido adiposo.
- *Heterogeneously Dense*: Senos con una mezcla de tejido denso y adiposo.
- *Scattered Fibroglandular*: Senos con una distribución irregular de densidades DIDSr, 2024c.

Simulación de Compresión Mamaria

Cada *phantom* fue sometido a una simulación de compresión utilizando el software VIC-TRE. Esta simulación replicó el proceso de compresión que ocurre durante la adquisición de mamografías, permitiendo modelar con precisión la deformación del tejido mamario DIDSr, 2024a.

Ground Truth

Además de las mamografías, se generaron datos de *Ground Truth*, que incluían mapas de desplazamiento y características de intensidad. Estos datos fueron esenciales para evaluar la precisión del algoritmo en la detección del *Air gap* y la corrección de las variaciones de intensidad, proporcionando una referencia objetiva con la cual comparar los resultados.

Apéndice 2. Algoritmo Propuesto

El algoritmo desarrollado en este estudio está diseñado para corregir las variaciones de intensidad en mamografías digitales causadas por la deformación del seno durante la compresión. El proceso seguido por el algoritmo se describe en detalle a continuación:

Detección del Air Gap

El primer paso del algoritmo consiste en la detección automática del *Air gap* entre las paletas de compresión y el seno deformado. Esta detección se realiza directamente en la imagen mamográfica 2D, sin necesidad de información adicional sobre el proceso de adquisición.

Aplicación del Modelo de Corrección Exponencial

Una vez detectado el *Air gap*, se aplica un modelo de corrección exponencial para ajustar las intensidades en la imagen. Este modelo está diseñado para compensar las variaciones de intensidad causadas por la deformación del tejido mamario, mejorando así la uniformidad y claridad de la imagen resultante.

Evaluación y Comparación

Finalmente, se evalúa la efectividad de la corrección comparando las imágenes corregidas con imágenes de referencia generadas a partir de datos de *Ground Truth*. Se utilizan métricas como el índice de similitud estructural (SSIM) y el error cuadrático medio (RMSE) para cuantificar la mejora en la calidad de la imagen.

