

Design guides for the development of a system to create interpretations of breast cancer studies

Francisco Eduardo Martínez-Pérez, Rosario Margot Camargo-Zebadúa, Sandra Edith Nava-Muñoz, Alberto Salvador Núñez-Varela, Abel de Jesús Guerrero-Jaime, José Ignacio Núñez-Varela, Francisco Edgar Castillo-Barrera, César A. Ramírez-Gómez

Published: 30 November 2024

Abstract

Breast cancer is a disease that, in the last years, has increased considerably worldwide. From a medical point of view, there exists a well-defined classification that radiologists use to interpret mammograms. New computational tools could be created with the purpose of helping the radiologists, to achieve this it is necessary to transform such medical knowledge into computational knowledge. This way, these computational tools could be used to help making decisions and accelerate the interpretations that radiologists perform. This research work presents design guides, which should be considered for the creation of computational systems. These guides are based on the knowledge that has been gathered from a breast cancer project in a hospital at the city of San Luis Potosí.

Keywords:

Breast cancer, design guides, case study, computational systems.

1 Introducción

En la actualidad, una de las enfermedades con mayor prevalencia en el mundo es el cáncer de mama. Con base en datos la Organización Mundial de la Salud, en 2022, en todo el mundo se diagnosticaron 2,3 millones de casos de cáncer de mama en mujeres, y se registraron 670 000 defunciones por esa enfermedad [1]. Desde el punto de vista médico, existe una metodología y guías que permiten realizar una clasificación del tipos de cáncer denominado Sistema de Informes y Datos de imágenes mamarias (por sus siglas en inglés Breast Imaging Reporting Data System

BIRADS) que pertenece al Colegio Americano de Radiología (por sus siglas en inglés American College of Radiology ACR).

El ACR ha establecido una serie de guías que deben ser consideradas para que un médico radiólogo pueda emitir una interpretación a partir de un estudio mamográfico. Un estudio mamográfico se compone de cuatro proyecciones tomadas a las mamas con cierta posición, denominadas cráneos caudales (CC, izquierda y derecha) y la proyección medio lateral oblicua (MLO, izquierda y derecha). En las proyecciones es posible encontrar hallazgos que están relacionados con lesiones y una referencia rápida de los hallazgos puede ser encontrada en [2]. En esta referencia rápida se hace mención de los siguientes hallazgos e información que se puede encontrar en un estudio mamográfico: i) Composición de la mama; ii) Masas; iii) Calcificaciones; y iv) Distorsión arquitectónica y en esta última hace mención a: a) Asimetrías; b) ganglio linfático intramamario; c) lesión de piel; d) conducto dilatado solitario; e) características asociadas y f) localización de la lesión. Todas estas son guías que presentan ciertas características específicas que deben ser consideradas para emitir una interpretación de un estudio mamográfico. Esta interpretación debe determinar la probabilidad de benignidad o malignidad del cáncer que se pudiese llegar a encontrar. Un médico radiólogo analiza todas las características anteriores para determinar esa probabilidad y emite la interpretación mencionando un tipo de composición de la mama y la clasificación BIRADS que puede ser 0, 1, 2, 3, 4a, 4b, 4c, 5 y 6, entre otros detalles. Es importante mencionar que todas las características antes mencionadas forman parte de la técnica de tamizaje primaria, que se realiza con el uso de estudios mamográficos. Dado que es una de las técnicas que se aborda en [2] y se complementan con el uso de ultrasonidos y de resonancias magnéticas según lo determine el radiólogo.

Con base en las características antes mencionadas, muchos sistemas computacionales han propuesto algunas herramientas [3] [4]. Sin embargo, se han limitado a herramientas aisladas que atacan una problemática computacional específica correspondiente a un tipo de hallazgo. Por lo que hasta el día de hoy no se ha encontrado una aplicación que considera todas las características que toma en cuenta un radiólogo [2][5] en un sistema computacional. Por lo que se han seguido guías de diseño abordando una característica a la vez, y no se han estructurado guías de diseño computacionales que abarquen toda la problemática del cáncer de mama.

Martínez-Pérez F. E., Nava-Muñoz S., Núñez-Varela, A., Núñez-Varela J. I., Castillo Barrera F. E., Ramírez-Gómez, C. A.
 Universidad Autónoma de San Luis Potosí, Facultad de Ingeniería.
 San Luis Potosí, S.L.P. México.
 {eduardo.perez, senavam, alberto.nunez, jose.nunez, ecastillo, crgomez}@uaslp.mx

Camargo-Zebadúa R. M., Guerrero-Jaime A. J.
 Hospital Central "Dr. Ignacio Morones Prieto"
 San Luis Potosí, S.L.P. México.
 {magocmd, abel.gro95}@gmail.com.

Adicionalmente, se pueden encontrar trabajos en los cuales utilizan diagramas de flujos o metodologías para el reconocimiento de calcificaciones y microcalcificaciones como las que se presentan en [4][3][6][7]. Mahesh et al., 2023 [4] proponen en su artículo el uso de Redes Neuronales convolucionales (CNN por sus siglas en inglés Convolutional Neural Network) con el objetivo de mejorar el diagnóstico de cáncer de mama. Estos autores utilizan bases de datos públicas en donde no describen las características que toman en cuenta de la referencia rápida que se presenta en [2], sino simplemente asumen que existen patrones en las imágenes de los datos públicos que es posible que no se repitan en otros casos. Balasubramaniam et al., 2023 [3], presentaron una de uso de la técnica de CNN en donde hacen alusión a los sistemas asistidos por computadora (por sus siglas en inglés Computer Aided Diagnosis CAD) como herramientas médicas. Ellos presentan algunos objetivos primarios de los sistemas de clasificación de cáncer de mama, sin embargo, no enfatizan las características primarias de los tipos de hallazgos y presentan con su técnica una clasificación benigna y maligna sin considerar las características específicas de esos patrones. Por último, en [6][7] se presenta metodologías que se basan en el reconocimiento de patrones en los cuales detectan calcificaciones y microcalcificaciones respectivamente. Sin embargo ambos trabajos no especifican a qué tipo de calcificaciones corresponden como lo son las de piel, las vasculares, las de palomitas de maíz entre otras; en donde debe existir una distinción entre macrocalcificaciones y microcalcificaciones como se definen en [8] para determinar una benignidad o malignidad.

Por lo que, en este artículo se presentan guías de diseño que permiten tomar en cuenta los pasos que deben ser considerados para la creación de un sistema de diagnóstico. Las guías de diseño se fundamentan en el estudio que se está realizando en el Hospital Central “Dr. Ignacio Morones Prieto” como estrategia en la creación de una metodología para la creación de interpretaciones de estudios mamográficos, en donde se comiencen a involucrar las herramientas computacionales basadas en las guías existentes para la valoración de cáncer de mama.

El resto del artículo está estructurado de la siguiente manera: En la sección 2 se presenta la metodología utilizada, posteriormente en la sección 3 se presenta la contribución del artículo que corresponde a las Guías de diseño. En la sección 4 se presenta la discusión con base en las guías de diseño propuestas, para finalizar con nuestras conclusiones en la sección 5.

2 Metodología

Actualmente, el proyecto de investigación se ha desarrollado en dos etapas. La primera etapa correspondió a un entendimiento de toda la problemática en la cual se invirtieron más de 600 horas de entrevistas, observaciones y reuniones de trabajo, que se realizaron a los expertos médicos, entre ellos radiólogos y oncólogos. A toda la información recopilada se ha aplicado la técnica de Teoría Fundamentada (TF) [9] como se podrá observar en las guías que se proponen. Con ésta se obtuvo un entendimiento más apegado a la forma en que los radiólogos realizan las interpretaciones de los estudios mamográficos. En esta etapa los radiólogos involucrados proporcionaron casos de estudios mamográficos. Estos fueron analizados por el equipo de trabajo. Posteriormente se realizaron reuniones de 2 a 3 horas semanales, por un periodo de 1 año, en las cuales iban describiendo caso por caso, en mamografías, los distintos hallazgos que se pueden encontrar con base en la distribución mostrada en [2].

La segunda etapa ha consistido en convertir el entendimiento de un radiólogo a un entendimiento computacional, en el cual ha sido necesario realizar adaptaciones por las guías proporcionadas en la referencia rápida de la ACR [2].

2.1 Motivación

En México, se han realizado campañas de toma de muestras de estudios mamográficos, como lo fueron los Convenios Específicos del Programa de Fortalecimiento a la Atención Médica [10]. Las campañas no incluyen profesionales para realizar diagnóstico o interpretación, como se muestran en los convenios de colaboración en materia de recursos presupuestarios federales con el carácter de subsidio para la operación del programa para el fortalecimiento a la atención médica. Inclusive dentro de las mismas capitales de los estados, en algunos hospitales, no se cuenta con médicos radiólogos y áreas con la tecnología que les permita realizar de una mejor manera las interpretaciones de los estudios mamográficos. Como sucede en el estado de San Luis Potosí, en el Hospital Central “Dr. Ignacio Morones Prieto” reciben estudios mamográficos de diversas instituciones de salud del estado para realizar las interpretaciones correspondientes, existiendo la necesidad de contar con más y mejores espacios para cubrir estas demandas en el menor tiempo posible.

2.2 Personas

En la primera etapa se formó un equipo de trabajo de profesionales médicos miembros del Hospital Central en el cual se involucraron: Un cirujano oncólogo con 10 años de experiencia y tres radiólogos con 12, 15 y 20 años de experiencia. En la segunda etapa, se formó el equipo con dos médicos radiólogos con 15 y 3 años de experiencia. Además de involucrar a un residente de la especialidad de radiología e imagen del Hospital Central.

2.3 Resultados

Como resultado de la primera etapa se utilizó la técnica de TF [9] a la información del entendimiento del cáncer de mama, cuyo análisis los realizó el equipo de trabajo con énfasis en desarrollo de sistemas para ir identificando y adaptando la información a un enfoque computacional. Como resultado de la TF se obtuvo una distribución de las calcificaciones y de las masas con base en la clasificación BIRADS, algunos resultados de reconocimiento de calcificaciones utilizando filtros de correlación se presentaron en [11]. En este mismo foro se presentó un modelo conceptual denominado PREVEMM (por sus siglas en inglés Pattern recognition to Evaluate MaMmograms).

Además, en la primera etapa, con base en la clasificación BIRADS, fue posible obtener la distribución de las masas, en donde se involucra las posibles combinaciones de su forma, el margen y la densidad.

Otro de los logros en la primera etapa, fue el conseguir cerca de 9000 imágenes que corresponden a poco más de 2100 estudios mamográficos para realizar nuestros análisis. Además se obtuvieron algunos conjunto de datos públicos (Dataset) como lo son INBreast [12], CBIS-DDSM [13], el BCDR [14] y MIAS [15]. Las imágenes obtenidas y las bases de datos públicas han permitido contrastar resultados e identificar oportunidades de investigación.

En la segunda etapa se ha estado obteniendo más información y diseñando el sistema, por lo cual en la siguiente sección se plantean las guías de diseño trabajadas hasta el momento.

3 Guías de Diseño para la creación de interpretaciones de cáncer de mama

El proceso de análisis de imágenes e información recolectada no ha sido un proceso sencillo. Como se mencionó anteriormente, fue posible obtener imágenes de distintos centros hospitalarios de la ciudad de San Luis Potosí, así como el descargar bases de datos públicas con el propósito de cruzar la información y observar las investigaciones realizadas en el estado del arte. Por lo anterior y con el objetivo de establecer la mejor solución en las interpretaciones de estudios mamográficos, se propone las siguientes guías de diseño:

1. Formato de entrada de un estudio mamográfico completo
2. Metodología de análisis de la problemática médica
3. Formato de salida con interpretación BIRADS
4. Análisis de las imágenes de estudios mamográficos
5. Procesamiento de imágenes para la creación de herramientas médicas

A continuación, se describen cada una de estas guías:

1. **Formato de entrada de un estudio mamográfico completo:** Con base en los estudios mamográficos proporcionados por los radiólogos y la TF se analizaron los tipos de imágenes, así como el número de imágenes que conforman un estudio mamográfico. Por lo que se consideran 4 imágenes por estudio que corresponden a 2 CC (Izquierda y Derecha) y 2 MLO (Izquierda y Derecha). Con base en las entrevistas con los expertos médicos, mencionaron que pueden existir tres tipos de situaciones que se pueden presentar con los pacientes cuando acuden al hospital a realizarse un estudio mamográfico que corresponde a i) Tamizaje; ii) Control; y iii) Diagnóstico. Para los casos i) y iii) se obtendrán 4 imágenes por estudio, dado que el primer caso corresponde a revisiones de rutina cada uno o dos años. Y el iii) tercero corresponde cuando una persona se detectó alguna anomalía en su cuerpo, por lo cual acuden con un especialista. Para el caso ii) en algunas ocasiones pudieran ser únicamente 2 imágenes. Esto se debe a que ya existe el antecedente de un posible hallazgo positivo y el médico profesional únicamente solicita el estudio del lado afectado, por lo que se encontrarán 2 imágenes en ese estudio. Por ejemplo, de los resultados de reconocimientos de calcificaciones que se presentaron en [11], se trabajó con casos de estudios mamográficos completos (registrados con calcificaciones) que contenían 4 imágenes (63 casos completos); 6 casos con 3 imágenes; 546 casos que contenían 2 imágenes y 135 casos que únicamente contenían una imagen por paciente. Todos estos casos corresponden a la base de datos pública CBIS-DDSM [13].
2. **Metodología de análisis de la problemática médica:** Esta se logra al convertir el conocimiento médico a conocimiento computacional. Como se mencionó en la introducción, en un estudio mamográfico es posible encontrar hallazgos y otro tipo de situaciones en un paciente. El tamizaje, es una de las técnicas que ayudan a detectar el cáncer de mama en las etapas tempranas. Por esta razón se utiliza la información de [2] y en específico el uso de mamografías. Todas estas características se complementan con la descripción de ellas que puede ser encontrada en [5]. Los detalles mostrados en esta referencia pueden dar una mejor forma de pensar y trasladar el conocimiento médico a conocimiento computacional como se realizó en nuestro trabajo utilizando TF, a partir de realizar diversas categorías que abarcan las calcificaciones y masas. Por ejemplo, una forma común que realizan los médicos

radiólogos es la búsqueda de patrones utilizando una forma de zig-zag o en líneas, como recorrer la imagen de arriba hacia abajo recorriendo renglón por renglón de izquierda a derecha. Este proceso se repite para cada imagen del estudio mamográfico. En este proceso, el radiólogo involucra las propiedades y las dimensiones de masas y calcificaciones, con base en la distribución reconocida de estas.

3. **Formato de salida con interpretación BIRADS.** Con base en la guía de referencia [2], en las entrevistas al equipo de profesionales médicos se les preguntó:

¿cuáles son los pasos que se deben seguir para realizar una interpretación de un estudio mamográfico cuando se realiza un tamizaje con base en la clasificación BIRADS?

Se obtuvo lo siguiente:

Paso 1: Se obtiene la composición de la mama

Paso 2: Se observa en zig-zags tratando de reconocer Masas y Calcificaciones

Paso 2.1 Identificación de las características asociadas a las masas y calcificaciones

Paso 2.2 Localización de las lesiones de las masas y calcificaciones

Paso 3 Localización de asimetrías y distorsión arquitectural

Paso 4 Análisis del ganglio linfático intramamario

Paso 5 Verificación del ducto dilatado solitario.

Estos pasos son de suma importancia en la creación de las interpretaciones de los estudios de cáncer de mama, cada paso tiene sus características específicas que debe ser tomadas en cuenta. Por lo que a continuación se describen estos pasos, obtenidos por la TF, en los cuales se resalta los posibles categorías, propiedades y dimensiones relacionadas con los detalles y características que debiesen considerarse en la creación de este tipo de sistemas computacionales:

Paso 1. corresponde a la obtención de la composición de la mama. Este es un dato que está relacionado con las ediciones del BIRADS. Un buen trabajo se puede encontrar en [16], en el cual la autora presenta una actualización de la edición 4 (composiciones enumeradas números del 1 al 4) a la 5ª edición (Composiciones con letras de la A a la D). La principal característica de composición de las mamas es el conocer la densidad de la mama que está relacionada con los niveles de blancos en cada una de las imágenes. Es decir, una imagen con pocos niveles de blancos, en la 5ª edición, es categorizada como ACR "A" y corresponde a que los senos son casi enteramente grasos. Por lo que los hallazgos son fácilmente localizables. Por el otro lado, un ACR "D" significa que existen demasiados niveles de blancos y se ocultan las lesiones, por lo que se solicita algunos estudios adicionales o complementarios. En términos médicos, ACR "D", está definido como: los senos son extremadamente densos, lo que reduce la sensibilidad de la mamografía con base en [2]. En este sentido, las características médicas cambiaron de BIRADS 4.0 a BIRADS 5.0 en donde la edición 4.0 manejaba en la composición del tejido etiquetada como "1", en la cual la mama es casi toda ella grasa, por lo que existía un rango comprendido entre 0 y menor que 25 % de los colores de blanco. Esta composición se convierte en la 5ª edición en ACR "A". Cada una de las cuatro composiciones maneja un rango de 25, siendo la última la 4ª clase que corresponde a la composición "D" con un rango mayor a 75% hasta 100% de blancos. La edición 4 facilitaba la interpretación de los valores blancos, sin embargo la 5ª

edición se vuelve subjetiva al suponer un aumento de la variabilidad inter e intraobservador como lo menciona Torres en [16]. Por lo que habrá que validar con los médicos radiólogos.

Paso 2. Se observan en zig-zags cada una de las dos lesiones más importantes que corresponden a Masas y Calcificaciones con base en [2].

- Las masas cuentan con una clasificación en dos niveles adicionales. El primer nivel lo conforman la Forma, los Márgenes y la Densidad. El segundo nivel está constituido por las diversas formas que se pueden localizar. Así como los diferentes tipos de márgenes y la densidad de las masas. Siendo esta una categoría definida con la TF.

Para este tipo de clases, en nuestra investigación, ya se tienen definidas las combinaciones que se pueden llegar a presentarse y fueron obtenidas por los radiólogos en la primera etapa de nuestro estudio.

- Las calcificaciones también cuentan con una clasificación de dos niveles que es considerada en los datos obtenidos por la TF. Sin embargo, antes de continuar es necesario considerar el tamaño de las calcificaciones como se propone en [8], en donde las macrocalcificaciones son consideradas aquellas que son mayor o igual a 0.5 mm y las microcalcificaciones de tamaño menor que 0.5 mm. De aquí que es necesario considerar los factores de conversión obtenidos en la realización del estudio mamográfico y almacenados en los archivos con extensión DCM. Con base en nuestro estudio de entendimiento, las calcificaciones debiesen ser consideradas en tres niveles de clasificación como lo presentamos en [11], en donde se agrega el tipo de BIRADS que corresponde con cada clase que se pueden encontrar.

En este paso, como resultado serán características específicas de los hallazgos. Estos deberán de ser conjuntados para encontrar las relaciones que existen entre ellos y de esta manera emitir una decisión lo más apegada a un médico radiólogo.

Paso 3. Una asimetría se define como un trastorno del desarrollo de la mama. Por lo que deviene un desarrollo asimétrico de las mamas [17]. De aquí la necesidad de contar con un estudio mamográfico completo para contrastar las imágenes del estudio del lado derecho e izquierdo, con el propósito de comparar ambos lados e identificar asimetrías o distorsiones, para los casos de tamizaje y diagnóstico.

Paso 4. Ganglio linfático intramamario se puede considerar dentro de las mismas especificaciones al momento de reconocer las masas o las calcificaciones. Se puede observar algunas características en [18], en donde referencian a la forma, los márgenes, el tamaño, entre otras características. Es necesario solicitar una descripción más objetiva de un médico radiólogo de las características específicas para obtener una diferencia en el entendimiento. Este tipo de hallazgos suelen ser considerados no relevantes como lo mencionan en [16].

Paso 5. Ducto dilatado solitario. Este tipo de hallazgos es considerado raro como se describe en [19]. El hallazgo es un conducto mamario asimétricamente dilatado que tiene un diámetro mayor a 2 mm o una porción ampular de más de 3 mm de diámetro [20]. Al ser un ducto tiene similitud con los tipos de calcificaciones vasculares. Se involucra un diámetro, por lo que es necesario considerar la información del archivo DCM, en la cual involucra un factor número para determinar las distancias en milímetros. Este tipo de hallazgos puede ser reconocido al momento de aplicar las técnicas en las masas o en las calcificaciones. Sin embargo, es necesario conocer las características específicas avaladas por un profesional médico.

4. **Análisis de las imágenes de estudios mamográficos.** Es importante tener conocimiento de los archivos que produce un mastógrafo por lo que se debe considerar lo siguiente:
 - a. Se debe realizar un entendimiento del tipo de imágenes que produce un mastógrafo. Una referencia del análisis de ellos se puede encontrar en [21]. Este autor presenta un análisis de los tipos de mastógrafos existentes en el mercado. Un detalle importante son los bits de resolución que se muestran en el artículo que corresponden a 13 y 14. Este factor es importante, dado que es la escala de grises de 13 bits, que corresponden a un valor entre 0 y 8,194 y la escala de grises de 14 bits corresponde a un valor entre 0 y 16,383. La resolución sirve para resaltar detalles de las lesiones tales como las microcalcificaciones. Estos tipos de mastógrafos producen imágenes de alta calidad. Por lo que entre mayor sea el número de bits, mayor será el detalle que se puede alcanzar y resaltar.
 - b. Los tipos de imágenes que produce un mastógrafo son del tipo DICOM® [22], con extensión DCM. Este tipo de archivos no es un archivo común, dado que representa los metadatos de las características con las que se toman las mamografías. Este tipo de archivos almacena la información del paciente, las características utilizadas en la toma de las imágenes, un arreglo (una imagen) o arreglos de imágenes según los modelos, entre otros datos. El arreglo de cada imagen varía en tamaño y pueden ser de 2736 x 3580, 3328 x 4084 o 2560 x 3328 píxeles entre otros tamaños similares. Los ajustes que realiza el técnico radiólogo, brillo y contraste entre otros, son registrados en el archivo de metadatos de cada imagen. Por lo que es de suma importancia mantener este tipo de información en cada imagen del estudio.
5. **Técnica para la creación de herramientas médicas:** Se realizó también un estudio al estado del arte en el cual se aplicó TF. Se identificó que se utilizan diversas técnicas de procesamiento de imágenes en los reconocimientos de hallazgos que se lleva en las guías para la interpretación del cáncer de mama como se abordaron en la introducción. La estructura más utilizada es con el uso de Reconocimiento de Patrones [23] en su diversas modalidades. Este cuenta con tres etapas para cada imagen que corresponden a:
 - a. Preprocesamiento en la cual se sugiere aplicar diversos procesos en las imágenes tales como: reducir el ruido; mejora del contraste o el rango dinámico de la imagen; mejorar la información de

los bordes en la imagen; analizar la imagen; y otros procesos específicos de cada aplicación. En este sentido algunos de estos procesos han sido implementados en nuestros trabajos como los presentados en [11].

- b. Extracción de las características: El propósito de encontrar las características es producir pocos descriptores que capturen la esencia de la imagen entrante. Es decir, identificar si existen objetos que son de interés para el estudio. Por ejemplo, para el caso de cáncer de mama, es necesario saber si existen calcificaciones y masas. No obstante, los valores de estos hallazgos pueden estar en diferentes rangos de la escala de grises que se maneja.
- c. Clasificación: El propósito de esta etapa es asignar una clase a las características identificadas en la etapa anterior. Los clasificadores son diseñados para optimizar algunas métricas tales como la probabilidad del error de clasificación o contar el error empírico. En este sentido, la clasificación BIRADS cuenta con 9 clases que dependerán de la primera validación del tipo de imagen. Es decir, primero es necesario realizar la clasificación de la composición de la mama, como se mencionó en la introducción, que podría ser del tipo ACR A, B, C o D. Como se pudo observar en el Paso 1 de la guía número 3 de este documento.

4 Discusión

En esta sección muestra una pequeña discusión con base en cada una de las guías antes expuestas que debiesen ser consideradas para la creación de sistemas de alta criticidad médica:

1. Todo sistema computacional enfocado en cáncer de mama o en situaciones médicas, debe de contar con un estudio completo si se desea aportar herramientas al estudio de tamizaje o *screening*. Si se utilizan bases de datos públicas, es necesario contar con los estudios completos avalados con las interpretaciones realizadas por médicos especialistas.
2. Para la identificación del cáncer de mama se tiene una metodología que los médicos radiólogos siguen para crear una interpretación de un estudio mamográfico. Sin embargo, es muy poco el trabajo que se presenta en el estado del arte en los cuales presenten características específicas de los hallazgos. Es decir, los investigadores toman las bases de datos públicas y se enfocan en encontrar patrones comunes, no obstante menciona una revista en [24] que dice:

"Desde la década de 1960, los radiólogos han notado que las mujeres tienen patrones únicos y ampliamente variables de tejido mamario visibles en la mamografía"

Por lo que se debe seguir la metodología médica y realizar las adaptaciones necesarias a un entendimiento computacional y no asumir existen patrones conocidos. El cáncer de mama debe ser considerado como un ambiente crítico en el estudio computacional.

3. Pocos son los sistemas computacionales que emiten una salida de sus resultados en la escala de BIRADS. La gran mayoría de los trabajos que se encuentran en el estado del arte se han limitado a entregar resultados como "Cáncer" o "No cáncer", o bien "Benigno" o "Maligno", siendo un resultado binario. Este tipo de sistemas utilizan bases de datos públicas, las cuales cuentan con una clasificación BIRADS, pero los autores no las consideran por el desconocimiento de la

problemática. Como las bases de datos públicas antes mencionadas.

4. Las imágenes que proporcionan los estudios mamográficos son de un tipo de datos especiales que buscan mostrar detalles muy finos, esto con base en el mastógrafo que se utiliza. Estos tienen resoluciones de 12, 13 o 14 bits. La gran mayoría de los trabajos que se pueden encontrar en el estado del arte manejan resoluciones máximas de 1024×1024 píxeles, aunque toman las imágenes fuente, el disminuir los tamaños de las imágenes implica pérdida de datos y por ende pérdida de características de algunos tipos de cáncer asociados a las microcalcificaciones e inclusive de las masas.
5. El reconocimiento de patrones es una técnica ampliamente utilizada en el estado del arte. Sin embargo, se ha estado utilizando indiscriminadamente los algoritmos de inteligencia artificial en donde se espera que lo resuelva todo sin proporcionar detalles de lo que se busca. Para el caso de situaciones médicas como el cáncer de mama se debe entender la problemática a resolver como un médico lo haría, pero en la realidad no sucede. Por lo que es necesario enfocar cada hallazgo con sus características y buscar nuevas maneras en las cuales se consideren las imágenes de su tamaño original. Dado que uno de los inconvenientes ha sido las limitaciones en el procesamiento computacional que tienen las herramientas actuales y que soportan como máximo de 1024×1024 píxeles, para poder ser procesadas por las redes neuronales artificiales.

5 Conclusiones

El campo de estudio del cáncer de mama para ayudar a la interpretación de estudios mamográficos presenta grandes retos y oportunidades. Una de las cosas que se debe de considerar es que el cáncer de mama es una enfermedad con criticidad alta. Por lo que es necesario buscar otras formas de llegar a los resultados que esperarían un médico radiólogo y no limitarse a un resultado binario y con pérdida de información necesaria para interpretar una mamografía por métodos computacionales.

El lograr muestras de cada una de las clases de BIRADS es un reto para el entrenamiento que se necesita en este tipo de enfermedades. Por lo que es necesario el apoyo de especialistas médicos en la creación de sistemas computacionales para la interpretación de cáncer de mama.

Este trabajo, con las guías de diseño propuestas, espera aportar en el desarrollo de sistemas que auxilien a los médicos en su trabajo, sin perder de vista los aspectos que deben ser indispensables y que han sido considerados en la medicina.

6 Acknowledgments

Se agradece al Comité Técnico y de Administración del Fideicomiso 23871 a través del Consejo Potosino de Ciencia y Tecnología (COPOCYT) del Gobierno del Estado de San Luis Potosí, que apoyaron al Proyecto titulado "Enlace Rosa: un entorno tecnológico para la valoración de cáncer de mama", así como al personal del Hospital Central "Dr. Ignacio Morones Prieto".

7 References

- [1] World Health Organization, "WHO Breast Cancer." Accessed: May 14, 2024. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/breast-cancer>
- [2] ACR, "ACR BI-RADS® Atlas Fifth Edition QUICK REFERENCE." Accessed: May 16, 2024. [Online].

- Available: <https://www.acr.org/-/media/ACR/Files/RADS/BI-RADS/BIRADS-Reference-Card.pdf>
- [3] S. Balasubramaniam, Y. Velmurugan, D. Jaganathan, and S. Dhanasekaran, "A Modified LeNet CNN for Breast Cancer Diagnosis in Ultrasound Images," *Diagnostics*, vol. 13, no. 17, pp. 1–28, 2023, doi: 10.3390/diagnostics13172746.
- [4] T. R. Mahesh *et al.*, "Transformative Breast Cancer Diagnosis using CNNs with Optimized ReduceLROnPlateau and Early Stopping Enhancements," *Int. J. Comput. Intell. Syst.*, vol. 17, no. 1, 2024, doi: 10.1007/s44196-023-00397-1.
- [5] D. M. Ikeda and K. K. Miyake, *Breast Imaging: The Requisites*, 3rd Editio. Elsevier, 2017.
- [6] P. Shi, J. Zhong, A. Rampun, and H. Wang, "A hierarchical pipeline for breast boundary segmentation and calcification detection in mammograms," *Comput. Biol. Med.*, vol. 96, no. October 2017, pp. 178–188, 2018, doi: 10.1016/j.compbiomed.2018.03.011.
- [7] K. Loizidou, G. Skouroumouni, C. Nikolaou, and C. Pitris, "An Automated Breast Micro-Calcification Detection and Classification Technique Using Temporal Subtraction of Mammograms," *IEEE Access*, vol. 8, pp. 52785–52795, 2020, doi: 10.1109/ACCESS.2020.2980616.
- [8] "Understanding Breast Calcifications." Accessed: May 20, 2024. [Online]. Available: <https://www.breastcancer.org/screening-testing/mammograms/what-mammograms-show/calcifications>
- [9] A. Strauss and J. Corbin, *Basics of Qualitative Research: Techniques and procedures for developing grounded theory*. Thousand Oaks, CA, 1998.
- [10] S. S. A. y P. Gobierno de México, "Convenios Específicos del Programa de Fortalecimiento a la Atención Médica 2019 (antes Unidades Médicas Móviles)." Accessed: Aug. 16, 2024. [Online]. Available: <https://www.gob.mx/salud/acciones-y-programas/convenios-especificos-del-programa-de-fortalecimiento-a-la-atencion-medica-2019-antes-unidades-medicas-moviles>
- [11] Francisco E. Martínez-Perez *et al.*, "Using reference points for detection of calcifications in mammograms for medical active systems," in *Proceedings of the 15th International Conference on Ubiquitous Computing & Ambient Intelligence (UCAI 2023)*, Springer, Cham, 2023, pp. 39--50. doi: https://doi.org/10.1007/978-3-031-48306-6_4.
- [12] I. C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. J. Cardoso, and J. S. Cardoso, "INbreast: Toward a Full-field Digital Mammographic Database.," *Acad. Radiol.*, vol. 19, no. 2, pp. 236–248, 2012, doi: 10.1016/j.acra.2011.09.014.
- [13] M. Heath, K. Bowyer, D. Kopans, R. Moore, and P. Kegelmeyer Jr, "The digital database for screening mammography," in *Proceedings of the Fourth International Workshop on Digital Mammography*, 2000. doi: 10.1007/978-94-011-5318-8_75.
- [14] Ceta-ciemat, "BCDR." Accessed: Jul. 11, 2023. [Online]. Available: <https://www.ciemat.es/cargarAplicacionNoticias.do?jsessionid=70C792CCF8908DF0D7289D48D83DC7CB?identificador=391>
- [15] I. Suckling, J.; Parker, J.; Dance, D.; Astley, S.; Hutt, I.; Boggis, C.; Ricketts, "Mammographic Image Analysis Society (MIAS) Database v1.21 [Dataset];" Dataset; Digital Mammogram Database Exerpta Medica: Dordrecht, The Netherland. [Online]. Available: <https://www.repository.cam.ac.uk/handle/1810/250394>
- [16] M. Torres Tabanera, "Novedades de la 5.a edición del sistema breast imaging reporting and data system (BI-RADS®) del Colegio Americano de Radiología," *Rev. Senol. y Patol. Mamar.*, vol. 29, no. 1, pp. 32–39, 2016.
- [17] L. Tejerizo-López *et al.*, "Asimetría mamaria (anisomastia)," *Clin. Invest. Ginecol. Obstet.*, vol. 27, no. 1, 1999, [Online]. Available: <https://www.elsevier.es/es-revista-clinica-e-investigacion-ginecologia-obstetricia-7-articulo-asimetria-mamaria-anisomastia-8197>
- [18] A. G. V. Bitencourt *et al.*, "Intramammary lymph nodes: Normal and abnormal multimodality imaging features," *Br. J. Radiol.*, vol. 92, no. 1103, 2019, doi: 10.1259/bjr.20190517.
- [19] V. J. Ayres, L. C. Ramalho, S. C. L. Wludarski, and E. de F. C. Fleury, "Screening a solitary dilated duct in the breast: A pictorial essay," *Breast Cancer Targets Ther.*, vol. 13, pp. 505–512, 2021, doi: 10.2147/BCTT.S307842.
- [20] B. Panigrahi *et al.*, "Solitary Dilated Ducts Revisited: Malignancy Rate and Implications for Management," *Acad. Radiol.*, vol. 30, no. 5, pp. 807–813, 2023, doi: <https://doi.org/10.1016/j.acra.2022.08.018>.
- [21] G. A. Martínez-Chávez, "Equipo Médico para Clínicas de Mama una Mirada Desde la Práctica de la Ingeniería Clínica," *Rev. Mex. Ing. Biomédica*, vol. 40, no. 3, pp. 1–8, 2019.
- [22] D. and C. in M. Imaging, "DICOM." Accessed: Aug. 17, 2024. [Online]. Available: <https://www.dicomstandard.org/current>
- [23] B. V. K. Vijaya Kumar, A. Mahalanobis, and R. Juday, *Correlation Pattern Recognition*. Cambridge University Press Cambridge, England, 2005.
- [24] C. Crespo G, "Un método basado en inteligencia artificial detecta el cáncer del futuro," National Geographic.



© 2024 by the authors. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.