



UNIVERSIDAD PERUANA DE CIENCIAS APLICADAS

FACULTAD DE INGENIERÍA

PROGRAMA ACADÉMICO DE INGENIERÍA DE SISTEMAS

Modelos de Clasificación de Densidad Mamaria Utilizando Redes

Neuronales Convolucionales

TRABAJO DE INVESTIGACIÓN

Para optar el grado de bachiller en Ingeniería de Sistemas

AUTOR(ES)

Garrafa Nina, Sonia Barbara (0000-0001-8573-6380)

Olea Bernales, Zoila Lastenia (0000-0001-8887-0227)

ASESOR

Espejo Villaizan, Daniel Derek (0000-0002-6737-0837)

Lima, 21 de enero del 2023

DEDICATORIA

A cada uno de nuestros padres e hijo. Por el apoyo incondicional que nos han brindado a lo largo de este camino, para poder conseguir nuestro bachiller.

AGRADECIMIENTOS

En esta ocasión, damos gracias a Dios por que nos permitió tener la valiosa experiencia de poder realizar nuestros estudios en la Universidad Peruana de Ciencias Aplicadas – UPC, además por permitirnos ser profesionales en lo que nos apasiona.

Gracias a la formación recibida podemos dejar este proyecto de investigación como prueba, para conocimiento y desarrollo de las nuevas generaciones.

Finalmente, queremos agradecer a quienes revisan nuestro proyecto de investigación, y permitirnos que esta investigación sea parte del repertorio de información de la universidad.

RESUMEN

Para el desarrollo de la investigación nos basamos en dos puntos importantes. La primera es una previa planificación, la cual contuvo el planteamiento de las preguntas específicas referente al tema de investigación, luego se consideró las bases de datos que nos permitió obtener dichos documentos seleccionados correctamente. Además, se determinó las palabras claves que hicieron más precisa la búsqueda. Así también, se consideró los criterios de inclusión y exclusión y se diagramó el proceso para tener un panorama visual del flujo realizado.

El segundo punto fue la obtención de resultados y análisis, que luego de haber culminado la primera parte, en cada uno de los papers seleccionados, se realizó la revisión de los resultados teniendo en cuenta la técnica, el aporte que dejó, el proceso de validación utilizado y los resultados obtenidos. En la parte de análisis se realizó, valga la redundancia, el análisis cruzado considerando aquí los puntos que ha contenido cada pregunta, en este caso se refirió a los modelos, componentes, técnicas y validaciones utilizadas en la realización de la clasificación mamaria con redes neuronales convolucionales.

Finalmente, se realizó las conclusiones respectivas donde se mencionó que es necesario seguir ambos puntos, pero en base a las preguntas definidas, para que así se obtenga la información orientada al proyecto. Además, se evidenció, en cuanto a la información brindada por los investigadores, que a mayor cantidad de validaciones cruzadas mayor será la precisión que incluso pueda llegar a obtener un 99% en la clasificación de densidad mamaria.

Palabras clave: Densidad Mamaria; Aprendizaje de Transferencia; Redes Neuronales Convolucionales; Mamografías densas.

CONVOLUTIONAL NEURAL NETWORK FOR THE CLASSIFICATION OF BREAST DENSITY

ABSTRACT

For the development of the investigation we base ourselves on two important points. The first is a previous planning, which contained the formulation of the specific questions regarding the research topic, then the databases that allowed us to obtain these correctly selected documents were considered. In addition, the keywords that made the search more precise were determined. Likewise, the inclusion and exclusion criteria were considered and the process was diagrammed to have a visual panorama of the flow carried out.

The second point was the obtaining of results and analysis, which after having completed the first part, in each of the selected papers, the review of the results was carried out taking into account the technique, the contribution it left, the validation process used and the results obtained. In the analysis part, the crossover analysis was carried out, considering here the points that each question has contained, in this case it referred to the models, components, techniques and validations used in carrying out the mammary classification with neural networks. convolutionals.

Finally, the respective conclusions were made where it was mentioned that it is necessary to follow both points, but based on the defined questions, so that the project-oriented information is obtained. In addition, it was evidenced, in terms of the information provided by the researchers, that the greater the number of cross-validations, the greater the precision that even 99% can obtain in the classification of breast density.

Keywords: Breast Density; Transfer Learning; Convolutional Neural Networks; Dense Mammography.

N° 3716_Modelos de Clasificación de Densidad Mamaria Utilizando Redes Neuronales Convolucionales

INFORME DE ORIGINALIDAD

4%	3%	2%	0%
INDICE DE SIMILITUD	FUENTES DE INTERNET	PUBLICACIONES	TRABAJOS DEL ESTUDIANTE

FUENTES PRIMARIAS

1	hdl.handle.net Fuente de Internet	1%
2	K. Pesce, M. Tajerian, M.J. Chico, M.P. Swiecicki, B. Boietti, M.J. Frangella, S. Benitez. "Estudio de la variabilidad inter- e intraobservador en la determinación de la densidad mamaria según la 5.a edición del Atlas BI-RADS®", Radiología, 2020 Publicación	1%
3	ccc.inaoep.mx Fuente de Internet	<1%
4	www.ic.uma.es Fuente de Internet	<1%
5	repositorio.utn.edu.ec Fuente de Internet	<1%
6	www.coursehero.com Fuente de Internet	<1%
7	repositorio.ulasamericas.edu.pe Fuente de Internet	

		<1 %
8	Anguita, M.. "Utility of NT-proBNP for Diagnosing Heart Failure in a Heterogeneous Population of Patients With Dyspnea. Spanish Multicenter Study", Revista Espanola de Cardiologia (Internet), 200605 Publicación	<1 %
9	repository.unab.edu.co Fuente de Internet	<1 %
10	revistascientificas.cuc.edu.co Fuente de Internet	<1 %
11	upc.aws.openrepository.com Fuente de Internet	<1 %
12	www.nature.com Fuente de Internet	<1 %

Excluir citas

Apagado

Excluir coincidencias < 20 words

Excluir bibliografía

Activo

TABLA DE CONTENIDOS

1	PLANIFICACIÓN	1
1.1	PLANIFICACIÓN.....	¡ERROR! MARCADOR NO DEFINIDO.
1.1.1	Planteamiento de preguntas de investigación..	¡Error! Marcador no definido.
1.1.2	Base de Datos	1
1.1.3	Palabras claves	1
1.1.4	Criterios de inclusión y exclusión	2
1.1.5	Diagrama del proceso para la revisión.....	2
2	RESULTADOS Y ANÁLISIS.....	7
2.1	RESULTADOS.....	7
2.1.1	Artículos.....	8
2.2	ANÁLISIS	32
2.2.1	Modelos (RQ1).....	33
2.2.2	Componentes (RQ2)	34
2.2.3	Técnicas (RQ3).....	36
2.2.4	Validación (RQ4)	38
2.2.5	Análisis cruzado	40
3	CONCLUSIONES.....	42
4	REFERENCIAS	43

ÍNDICE DE TABLAS

Tabla 1	2
Tabla 2	3
Tabla 3	7
Tabla 4	32
Tabla 5	33
Tabla 6	34
Tabla 7	36
Tabla 8	38
Tabla 9	40
Tabla 10	40

ÍNDICE DE FIGURAS

Figura 1.	2
Figura 2.	7
Figura 3.	8
Figura 4	27
Figura 5	32

INTRODUCCIÓN

El presente documento está orientado a brindarnos información detallada de otras investigaciones que nos sirvan como punto de partida o como referencia para una nueva comprobación acerca de la clasificación de las mamas densas con redes neuronales convolucionales. La densidad mamaria, actualmente se relaciona con el aumento del riesgo de cáncer de mama, ya que la imprecisión de la clasificación en la mamografía, dificulta la detección del cáncer. Por consiguiente, se realiza la búsqueda respecto a la problemática en los diferentes metabuscadores (Science Direct, Scopus e IEEE) para obtener las mejores bases de información para el desarrollo de nuestro proyecto.

Para mitigar este problema nos basaremos en cuatro preguntas fundamentales que nos permitan obtener herramientas con las que se pueda atacar el problema y plantear la mejora de la problemática.

Por último, se identificará las palabras claves, criterios de inclusión y exclusión, así como también se definirá el proceso que se ha seguido para dicha revisión, luego vendrán los resultados y el análisis para que podamos emitir nuestras propias conclusiones acerca de esta revisión.

1 PLANIFICACIÓN

El desarrollo del estado del arte se realizará aplicando la revisión sistemática de la literatura que están determinados por ciertos lineamientos. (Kitchenham, 2007; Wong, 2017).

Las fases que se usarán son las siguientes: (1) Planificación de la revisión, (2) Desarrollo de la revisión y, (3) Resultados y análisis. [En esta sección obligatoria se inicia el contenido o cuerpo del documento.

1.1 Planteamiento de preguntas de investigación

Para la planificación realizada en la presente investigación se consideró las siguientes preguntas:

PI1: ¿Qué modelos de clasificación de densidad mamaria existen?

PI 2: ¿Qué componentes tienen los modelos de clasificación de densidad mamaria?

PI 3: ¿Cuáles son las técnicas que se pueden usar para construir estos componentes de clasificación?

PI 4: ¿Cómo se valida los modelos de clasificación de densidad mamaria?

1.2 Base de Datos

Se realizó la búsqueda de los papers en los siguientes metabuscadores, de los cuales se obtuvo los 20 papers elegidos.

- IEEE
- ScienceDirect
- Scopus

1.3 Palabras claves

Las palabras claves consideradas para la búsqueda fueron, en especial, en inglés debido a que existe mayor información e investigación en papers de procedencia extranjera.

- Inteligencia artificial
- Breast density classification
- Convolutional neural networks
- Breast cancer
- Mammogram with breast density

1.4 Criterios de inclusión y exclusión

Tabla 1

Criterios de inclusión y exclusión.

Criterios de inclusión	Criterios de exclusión
Sistemas de clasificación de densidad mamaria con mamografías.	Sistemas de clasificación de densidad mamaria con ecografías o tomografías.
Papers relacionados a la clasificación de la densidad mamaria con el uso de la técnica inteligencia artificial.	Papers relacionados a la clasificación de densidad mamaria realizadas manualmente.
Papers que se encuentren en el rango de 5 años. (desde la fecha hacia abajo).	Papers con idiomas diferentes al inglés o español.
Se revisará papers que estén catalogados como artículos de investigación.	No se considerará en la búsqueda artículos de revisión.

1.5 Diagrama del proceso para la revisión

Figura 1

Diagrama del proceso para la revisión sistemática de la literatura.

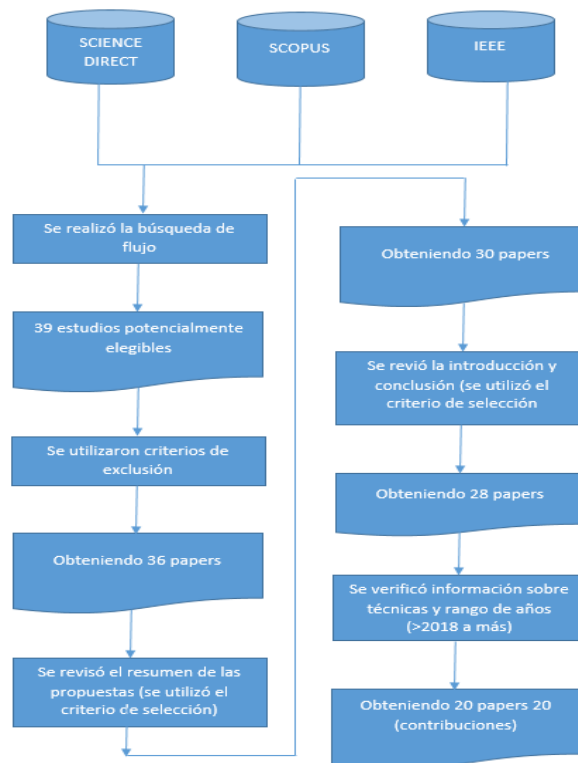


Tabla 2*Lista de los artículos seleccionados*

ID.	Título	Fuente	Pregunta de investigación	Cuartil	Metabuscador
P01	Uso de una red neuronal convolucional basada en ondas y ajustada para la clasificación de la densidad mamaria en imágenes mamográficas	Matsuyama et al., (2020)	Q3	RQ3	ScienceDirect
P02	Clasificación de las categorías de densidad mamaria según las redes neuronales SE-Attention.	Deng et al., (2019)	Q1 Q2	RQ1 RQ2	ScienceDirect
P03	Redes neuronales convolucionales residuales para extraer automáticamente características de densidad mamaria significativas.	Lizzi et al., (2019)	Q3	RQ3	ScienceDirect
P04	Redes neuronales convolucionales de múltiples vistas para la clasificación de imágenes mamográficas	Sun et al., (2019)	Q3	RQ3	IEEE
P05	Una nueva arquitectura de aprendizaje profundo que supera en rendimiento el aprendizaje por transferencia "estándar" y los métodos basados en funciones en la evaluación automatizada de la densidad mamográfica	Trivizakis et al., (2019)	Q3	RQ3	Scopus

P06	Internet de las cosas médicas que integra el aprendizaje profundo con datos aumento para clasificación de densidad de mamografía	Sadad et al., (2021)	Q3	RQ3	Scopus
P07	Un sistema de aprendizaje profundo para obtener los parámetros óptimos para un segmentación de tejido denso y mama basada en umbrales	Pérez-Benito et, al. (2020)	Q3	RQ3	ScienceDirect
P08	Clasificación de densidad mamográfica de múltiples vistas por aprendizaje residual dilatado y guiado por la atención	Li et al., (2020)	Q3	RQ3	IEEE
P09	Validación externa de un profundo Modelo de aprendizaje para predecir Densidad mamaria mamográfica en Práctica clínica de rutina	Dontchos et al., (2020)	Q4	RQ4	ScienceDirect
P10	Aprendizaje federado para la clasificación de la densidad mamaria: una implementación en el mundo real	Roth et al., (2020)	Q1 Q2	RQ1 RQ2	IEEE
P11	Deep-LIBRA: un método de inteligencia artificial para la cuantificación sólida de la densidad mamaria con validación independiente en	Maghsoudi et al., (2021)	Q1 Q2	RQ1 RQ2	ScienceDirect

	la evaluación del riesgo de cáncer de mama				
P12	Evaluación multiinstitucional y evaluación de crowdsourcing del aprendizaje profundo para la clasificación automatizada de la densidad mamaria	Chang et al. (2020)	Q3	RQ3	ScienceDirect
P13	Detección y clasificación simultánea de masas mamarias en mamografías digitales a través de un sistema CAD de aprendizaje profundo basado en YOLO	Al-masni et al. (2018)	Q1 Q2	RQ1 RQ2	ScienceDirect
P14	Artemisia: validación de un modelo de aprendizaje profundo para la categorización automática de la densidad mamaria	Tajerian et al., (2021)	Q4	RQ4	Scopus
P15	Aprendizaje profundo en segmentación y clasificación de imágenes de mamografía: enfoque de CNN automatizado	Salama (2021)	Q3	RQ3	ScienceDirect
P16	Clasificación de densidad mamaria en mamografías: una investigación de técnicas de codificación en patrones locales basados en binarios	Rampun et al., (2020)	Q3	RQ3	ScienceDirect

P17	Segmentación y clasificación de densidad mamaria totalmente automatizada mediante aprendizaje	Saffari et al., (2020)	Q3	RQ3	Scopus
P18	Redes neuronales convolucionales residuales para la clasificación de densidad mamaria	Lizzi et al., (2019)	Q3	RQ3	ScienceDirect
P19	Validación de un nuevo software totalmente automatizado para la evaluación de la densidad mamaria mamográfica digital 2D para predecir el riesgo de cáncer de mama	Gorgi et al., (2021)	Q4	RQ4	Scopus
P20	Aprendizaje profundo para la evaluación de la densidad mamaria mamografía y más allá	Chan (2019)	Q3	RQ3	Scopus

2 RESULTADOS Y ANÁLISIS

2.1 Resultados

Luego de determinar las palabras claves a considerar y los portales de búsqueda a utilizar, se obtuvo más de 39 resultados que, a través de los criterios de selección, la restricción de fecha de publicación y una exhaustiva revisión de los resúmenes y resultados de las investigaciones, se convirtieron en un total de 20 artículos seleccionados.

Tabla 3

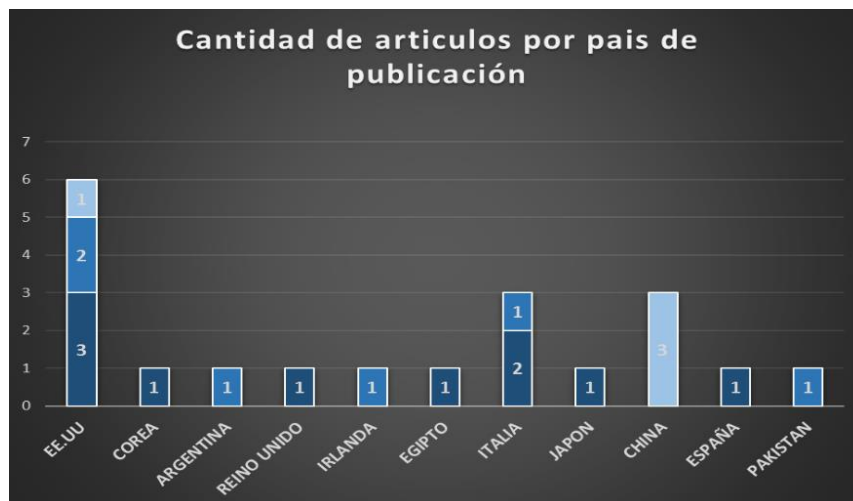
Cantidad de Papers Encontrados vs Seleccionados

Metabuscadores	Papers Encontrados	Papers Seleccionados
ScienceDirect	18	10
Scopus	11	6
IEEE	10	4
Total	39	20

En el siguiente gráfico se observa los países en los que se llevó a cabo las investigaciones. De ellos se puede mencionar que el 30% fueron de EEUU, el 26% se realizó en ITALIA y CHINA y el resto de los países representa en total el 44%.

Figura 2

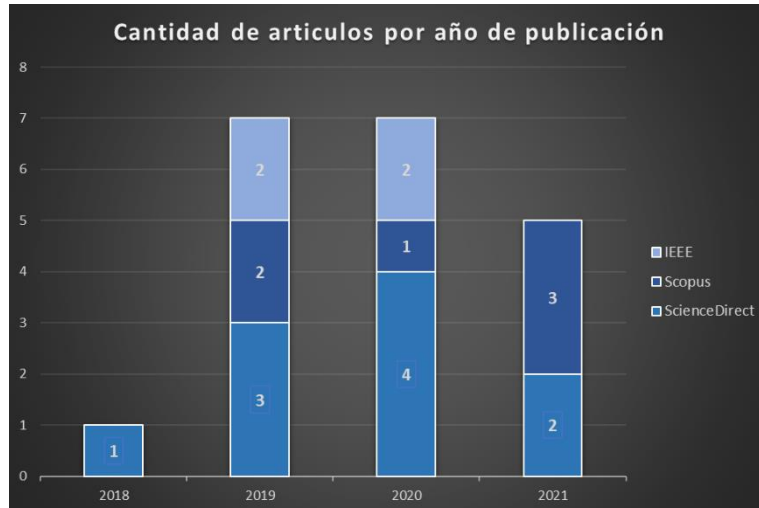
Cantidad de artículos por país de publicación



Asimismo, en este gráfico se encuentra que el 65% de las publicaciones son entre los años 2020 y 2019, el 30% pertenece al año 2021 y el otro 5% fueron realizados en el año 2018.

Figura 3

Cantidad de artículos por año de publicación



2.1.1 Artículos

2.1.1.1 Artículo 01 (Matsuyama et al., 2020):

Título: Uso de una red neuronal convolucional basada en ondas y ajustada para la clasificación de la densidad mamaria en imágenes mamográficas

Aporte:

El propósito de este estudio fue desarrollar una herramienta computarizada útil para ayudar a los radiólogos a determinar la categoría de densidad mamaria de la paciente en la mamografía, donde se presenta un modelo para clasificar automáticamente las densidades mamarias mediante el empleo de una red neuronal convolucional (CNN) basada en transformadas de ondas y sintonizada. A diferencia de los métodos comunes basados en CNN que utilizan imágenes originales o preprocesadas como entradas, adoptaron el uso de coeficientes wavelet redundantes en el nivel 1 como entradas para el modelo CNN. El estudio se centró principalmente en discriminar entre densidad dispersa y densidad heterogénea, que son las dos categorías de densidad más difíciles de diferenciar para los radiólogos. El sistema propuesto logró un 88,3% de precisión general. Para demostrar la eficacia y utilidad del método propuesto, los resultados obtenidos de un modelo CNN de ajuste fino convencional se compararon con los del método propuesto.

Proceso o validación:

Utilizando AlexNet, que es un modelo de CNN muy conocido para la clasificación de la densidad mamaria. AlexNet, que ha sido previamente entrenado con ImageNet, consta de cinco capas convolucionales y tres capas completamente conectadas (FC). Se construyó una nueva red de CNN utilizando las capas anteriores de AlexNet previamente entrenado. El conjunto de datos utilizado, la extracción de la información espectral de la imagen mediante transformadas de ondículas y la arquitectura del modelo CNN afinado.

Resultado:

Los resultados demuestran que la técnica propuesta es muy prometedora para ayudar a los radiólogos y servirles como un segundo ojo para clasificar las categorías de densidad mamaria en la detección del cáncer de mama. Los resultados obtenidos de un modelo CNN de ajuste fino convencional se compararon con los del método propuesto. Los resultados demuestran que la técnica propuesta es muy prometedora para ayudar a los radiólogos y servirles como un segundo ojo para clasificar las categorías de densidad mamaria en la detección del cáncer de mama. Los resultados obtenidos de un modelo CNN de ajuste fino convencional se compararon con los del método propuesto. Los resultados demuestran que la técnica propuesta es muy prometedora para ayudar a los radiólogos y servirles como un segundo ojo para clasificar las categorías de densidad mamaria en la detección del cáncer de mama. La precisión general alcanzó el 88,3% para el método.

2.1.1.2 Artículo 02 (Deng et al., 2019):

Título: Clasificación de las categorías de densidad mamaria según las redes neuronales SE-Attention

Aporte:

Se propone un marco de red neuronal convolucional (CNN) mejorado que integra un mecanismo innovador de atención SE para aprender características discriminatorias, con el objetivo de la clasificación automática de BD en mamografía. Se construyó un nuevo conjunto de datos de evaluación comparativa a partir de 18157 imágenes de BD, segmentadas manualmente en 4 niveles según el Sistema de datos de informes e imágenes de mama (BI-RADS): A (graso), B (fibroglandular), C (heterogéneamente denso) y D (extremadamente denso). El método propuesto consta de tres fases principales: (i) mejora de datos y normalización de imágenes de mama (ii) entrenamiento SE-Atención para la

recalibración y fusión de características para clasificar mejor la densidad y (iii) diseño de la pérdida auxiliar. Se adopto un enfoque de atención donde se utiliza el mecanismo SE-Attention para aprender las características de densidad.

Proceso o validación:

El modelo de aprendizaje profundo se compone de ANN multicapa, que es la clave para realizar la inteligencia artificial (IA). Es decir, implementa la aproximación de un algoritmo complejo no lineal de alta dimensión, caracterización de los datos mediante el aprendizaje de estructuras no lineales. CNN es uno de los modelos de aprendizaje profundo más comunes. En comparación con el aprendizaje automático tradicional, puede evitar el complejo proceso de extracción de características. El rendimiento de Inception-V4-SE-Attention se analiza a partir de la matriz de confusión del modelo en el conjunto de datos de prueba.

Resultado:

Los resultados experimentales demuestran que el marco propuesto obtiene una mayor precisión de clasificación que la red original, como Inception-V4, ResNeXt, DenseNet, aumentando el rendimiento del 89,97% al 92,17%, del 89,64% al 91,57%, del 89,20% al 91,79% respectivamente. Entre ellos, Inception-V4 mejorado posee la mayor precisión, mientras que DenseNet mejora en la mayor medida, tanto el método original como el mejorado.

2.1.1.3 Artículo 03 (Lizzi et al., 2019):

Título: Redes neuronales convolucionales residuales para extraer automáticamente características de densidad mamaria significativas

Aporte:

Se realizo la clasificación de densidad mamaria con una red neuronal residual profunda y los análisis futuros que podría realizar. La densidad mamaria es una de los más importantes factores de riesgo de cáncer de mama y representa la cantidad de fibroglandular tejido con respecto al tejido graso. Se entreno una CNN profunda realizar la clasificación de la densidad mamaria de dos formas. Primero, se clasifico mamografías utilizando dos "superclases" que son densas y no densas seno. En segundo lugar, se entrenó la red neuronal residual para clasificar las mamografías. Según las cuatro clases del estándar BI-RADS.

Proceso o validación:

En el estándar BI-RADS, la discriminación entre mama densa y no densa significa clasificar dos "superclases", la hecha de mamografías pertenecientes a las clases A y B y el otro hecho a las clases C y D. Este problema tiene una relevancia clínica ya que una mujer con una mama densa debe ser examinada más con cuidado. El conjunto de datos AOUP se ha dividido aleatoriamente en el conjunto de entrenamiento (1356 exámenes), conjunto de validación (160 exámenes) y conjunto de prueba (160 exámenes). La clasificación las puntuaciones de las últimas capas de cada CNN se han promediado para producir una etiqueta que tiene en cuenta todas las imágenes relacionadas con un mismo tema. Es más, Se han explorado diferentes tamaños de imagen de entrada para comprender si existe una dependencia de las cifras de mérito del tamaño de entrada de la imagen. Por lo tanto, se entrenaron siete CNN diferentes por proyección con imágenes de 250×250 píxeles a 850×850 píxeles.

Resultado:

Respecto al problema denso / no denso, la CNN entrenó en 650×650 píxeles imágenes, predice la etiqueta correcta con una precisión igual al 89,4%, que es la mejor exactitud de la prueba obtenida en esta tarea. En cuanto a la clasificación BI-RADS, se obtuvo una exactitud de la prueba en imágenes de 650×650 píxeles igual al 77,3%.

2.1.1.4 Artículo 04 (Sun et al., 2019):

Título: Redes neuronales convolucionales de múltiples vistas para la clasificación de imágenes mamográficas

Aporte:

El aprendizaje profundo se ha aplicado ampliamente a la clasificación de imágenes mamográficas. Sin embargo, la mayoría de los métodos existentes se basan en una sola vista de mamografía y no pueden extraer suficientemente las características discriminatorias, lo que da como resultado una precisión de clasificación insatisfactoria. Se propuso una nueva red neuronal convolucional multivista basada en múltiples vistas mamográficas. Teniendo en cuenta que las imágenes adquiridas desde diferentes perspectivas contienen información de masa mamaria diferente y complementaria, se modificó la arquitectura de las redes neuronales convolucionales (CNN) para aprovechar la información complementaria de varias vistas de la mamografía. La nueva arquitectura puede extraer características

discriminatorias de las vistas oblicua medio lateral (MLO) y craneocaudal (CC) de las imágenes mamográficas y puede incorporar eficazmente estas características para las imágenes mamográficas.

Proceso o validación:

El método novedoso para la clasificación de imágenes mamográficas. El método propuesto integra MVCNN y MDCNN para extraer características de múltiples vistas de las imágenes de masa mamaria. La red puede extraer características con información complementaria y diversa, lo que es propicio para aumentar la precisión de clasificación y la robustez del modelo. Los resultados mostraron que el método superó los métodos comparados de vanguardia para la clasificación de imágenes mamográficas.

Resultado:

La base de datos digital para mamografías de detección (DDSM) es un conjunto de datos de imágenes mamográficas, que contiene aproximadamente 2600 imágenes médicas de mama de pacientes, cada paciente proporciona 4 imágenes de las vistas MLO y CC de las mamas izquierda y derecha respectivamente. Los conjuntos de datos de entrenamiento y prueba en los primeros seis pliegues contienen el 85% y el 15% de las muestras, respectivamente. En otras palabras, hay 1.228 muestras de entrenamiento y 217 muestras de prueba en los primeros seis pliegues. En el último pliegue, el 90% de las muestras (es decir, 1300 muestras) se seleccionan como muestras de entrenamiento y las muestras restantes (es decir, 145 muestras) se consideran muestras de prueba. El muestreo aleatorio de muestras se realiza en el subconjunto dividiendo las muestras en siete porciones por porcentajes sin superposición.

2.1.1.5 Artículo 05 (Trivizakis et al., 2019):

Título: Una nueva arquitectura de aprendizaje profundo que supera en rendimiento el aprendizaje por transferencia "estándar" y los métodos basados en funciones en la evaluación automatizada de la densidad mamográfica

Aporte:

En este estudio, se ha propuesto y evaluado metodologías avanzadas de aprendizaje automático con el objetivo de un método objetivo y confiable para la puntuación de la densidad mamaria a partir de imágenes mamográficas de rutina. La tubería de análisis de imágenes propuesta incorpora textura [filtros de Gabor y patrón binario local (LBP)] y

características basadas en gradientes [histograma de gradientes orientados (HOG) así como características robustas aceleradas (SURF)]. Además, también se utilizaron enfoques de aprendizaje por transferencia con pesos entrenados en ImageNet para la comparación, así como una red neuronal convolucional (CNN). El modelo CNN propuesto se entrenó completamente en dos conjuntos de datos de mamografía abiertos y se encontró que era la metodología de rendimiento óptimo (AUC hasta 87,3%).

Proceso o validación:

La arquitectura CNN 2D completamente entrenada consta de 15 capas, incluida la entrada de imagen de forma 725x234x1, 6 capas convolucionales cada una seguida de una capa de normalización por lotes, 2 capas completamente conectadas con 100 neuronas cada una y finalmente una capa de clasificación softmax, se seleccionó ReLU como la función de activación de las capas convolucionales con 30 a 240 granos por capa y un campo receptivo de 3x3. Además, se aplicó un 20% de probabilidad de abandono a las neuronas completamente conectadas. Se utilizó la metodología Glorot para la inicialización del peso. El código fuente completo y los hiperparámetros finales de la arquitectura 2D personalizada están disponibles en línea (<https://github.com/trivizakis/breast-density-analysis>). Optimización de hiperparámetros. El proceso de ajuste de una arquitectura profunda plantea una tarea desafiante de buscar los parámetros óptimos para descubrir el modelo con mejor rendimiento.

Resultado:

Todos los modelos se ajustaron en el mismo conjunto de convergencia estratificada (entrenamiento / validación) y se evaluaron en idénticos probar pliegues de validación cruzada para garantizar una comparación justa y transparente. Esto resultó en un 64,6% de formación, un 15,4% de validación y un 20% de pruebas de imágenes de mamografía del mini-MIAS y un 63,9% de formación, un 16,1% de validación y un 20% de pruebas de la base de datos DDSM respectivamente. Se realizaron diferentes algoritmos y estrategias de anotación en los dos conjuntos de datos estudiados para identificar la representación óptima del espacio de características. Las precisiones en el conjunto de datos mini- MIAS variaron del 50,9% (características seleccionadas de GABOR + LBP) al 74,2% (LBP) para la clasificación de tres clases, mientras que, para la clasificación binaria, las puntuaciones de AUC variaron del 48,7% (características seleccionadas de HOG) a 78,0% (LBP). De manera

similar, la metodología descrita anteriormente se aplicó en el conjunto de datos DDSM completo para predecir la puntuación MBD en un esquema de anotación binario y de clases múltiples. La puntuación óptima (ACC 79,3% y AUC 84,2%) para las técnicas basadas en características se observó en el análisis binario (mamografías densas o no densas) con el método SURF.

2.1.1.6 Artículo 06 (Sadad et al., 2021):

Título: Internet de las cosas médicas que integra el aprendizaje profundo con datos aumento para clasificación de densidad de mamografía

Aporte:

En esta investigación detecta BD automáticamente usando imágenes de mamografías basadas en Internet de Dispositivos compatibles con Medical Things (IoMT). Se aplicaron dos modelos de redes neuronales convolucionales profundas previamente entrenados llamados DenseNet201 y ResNet50 a través de un enfoque de aprendizaje por transferencia. Un total de 322 imágenes de mamografía que contienen 106 grasas, 112 densas y Se obtuvieron 104 casos glandulares del Conjunto de datos de la Sociedad de análisis de imágenes de mamografías. La poda de regiones irrelevantes y la mejora de las regiones objetivo se realizan en preprocesamiento. La precisión de clasificación general de la tarea BD se realiza y logrado 90.47% a través del modelo DensNet201. Tal marco es beneficioso en identificar BD más rápidamente para ayudar a los radiólogos y pacientes sin demora.

Proceso o validación:

Para la integración del diagnóstico asistido por computadora con la nube, se utiliza una máquina de mamografía basada en IoT para cargar las imágenes de mamografía en la nube. Se aplica un sistema CAD basado en la nube con un algoritmo avanzado de aprendizaje automático para categorizar los tejidos BD. Posteriormente, el informe de diagnóstico con respecto a las imágenes de mamografía se enviará de vuelta al iniciador, entregando así una respuesta oportuna a la paciente. El procedimiento de preprocesamiento se adopta normalmente en todas las aplicaciones de reconocimiento de patrones para mejorar la calidad de la imagen al eliminar el ruido y los detalles innecesarios para adquirir Regiones de interés (ROI) en imágenes digitales para lograr una alta precisión y reducir el costo computacional (Al-Ameen et al., 2015; Nodehi et al., 2014; Sadad et al., 2021). Además, el preprocesamiento mejora el proceso de entrenamiento del aprendizaje profundo porque

entrenar un modelo de CNN sin preprocesamiento podría degradar el rendimiento de la clasificación Amin, Sharif, Raza, Saba y Rehman, 2019. Por lo tanto, empleamos la ecualización de histograma adaptativo limitado por contraste (CLAHE) para la eliminación de ruido y El recorte manual se realiza para la selección de ROI.

Resultado:

El modelo previamente entrenado a través de ImageNet se utiliza para la clasificación de BD. El entrenamiento y las pruebas de los modelos se lograron utilizando el 80 y el 20% de las imágenes de mamografías en el conjunto de datos MIAS. Además, el proceso de barajado aleatorio se realiza durante el entrenamiento y las pruebas para lograr buenos resultados alcanzando la máxima precisión, puntuación de 91,3, 96 y 95,45% en los datos de prueba, respectivamente, para la clase glandular del conjunto de datos MIAS utilizando el modelo DensNet201. La precisión de clasificación general del modelo DenseNet20 y ResNet50 es 90,47 y 80,95%, respectivamente. Además, el valor de las estadísticas kappa también se calcula y logra 0.857 en la arquitectura DenseNet201. En la técnica propuesta, se emplean enfoques de aprendizaje profundo a través del aprendizaje por transferencia para la clasificación de BD en el conjunto de datos MIAS. El marco propuesto mejora la precisión de la clasificación en comparación con los métodos tradicionales de aprendizaje automático.

El modelo de aprendizaje profundo es eficiente en la clasificación porque no necesita extracción y selección de características. Simplemente podemos proporcionar imágenes como entrada al modelo en características artesanales proporcionadas a clasificadores de aprendizaje automático convencionales. El marco propuesto puede categorizar las imágenes de mamografía en clases densas.

2.1.1.7 Artículo 07 (Pérez-Benito et al., 2020):

Título: Un sistema de aprendizaje profundo para obtener los parámetros óptimos para una segmentación de tejido denso y mama basada en umbrales

Aporte:

La red sanitaria española estableció programas de cribado poblacional en todas las comunidades autónomas, donde se realizan mamografías a mujeres asintomáticas con fines de diagnóstico precoz. La densidad mamaria evaluada mediante mamografías digitales es un biomarcador conocido por estar relacionado con un mayor riesgo de desarrollar cáncer de mama. Por tanto, es fundamental proporcionar un método fiable para medir la densidad

mamaria a partir de mamografías. Además, la automatización completa de este proceso de segmentación se está volviendo fundamental a medida que aumenta la cantidad de mamografías cada día. Los desafíos importantes están relacionados con las diferencias en las imágenes de diferentes dispositivos y la falta de un estándar de oro objetivo. Este artículo presenta un marco totalmente automatizado basado en el aprendizaje profundo para estimar la densidad mamaria. El marco cubre la detección de mamas, la exclusión de los músculos pectorales y la segmentación del tejido fibroglandular.

Proceso o validación:

Un estudio multicéntrico, compuesto por 1785 mujeres cuyas mamografías "de presentación" fueron segmentadas por dos radiólogos experimentados. Un total de 4992 de las 6680 mamografías se utilizaron como corpus de entrenamiento y las restantes (1688) formaron el corpus de prueba. Este artículo presenta un paso de normalización del histograma que suavizó la diferencia entre la adquisición, una arquitectura de regresión que aprendió los parámetros de segmentación como características intrínsecas de la imagen y una función de pérdida basada en la puntuación DICE.

Resultado:

Los resultados obtenidos indican que el nivel de concordancia (puntuación DICE) alcanzado por los dos radiólogos (0,77) también fue alcanzado por el marco automatizado cuando se comparó con la segmentación mamaria más cercana de los radiólogos. Para los adquiridos con dispositivo de mayor calidad, la puntuación DICE por dispositivo de adquisición alcanzó 0,84, mientras que la concordancia entre radiólogos fue 0,76.

2.1.1.8 Artículo 08 (Li et al., 2020):

Título: Clasificación de densidad mamográfica de múltiples vistas por aprendizaje residual dilatado y guiado por la atención

Aporte:

En este estudio, presenta un enfoque radiómico basado en el aprendizaje residual dilatado y guiado por la atención para la tarea de clasificación de densidad mamográfica. El método propuesto se instancia con dos conjuntos de datos, un conjunto de datos clínicos y un conjunto de datos disponibles públicamente, y se obtuvieron precisiones de clasificación de 88,7% y 70,0%, respectivamente. Aunque la precisión de la clasificación del conjunto de

datos públicos fue menor que la del conjunto de datos clínicos, lo que probablemente estuvo relacionado con el tamaño del conjunto de datos, Nuestro modelo propuesto aún logró un mejor rendimiento que las redes residuales ingenuas y varios enfoques basados en el aprendizaje profundo publicados recientemente. Además, diseñamos una arquitectura de red multiseuencia dirigida específicamente al análisis de las mamografías multivista. Utilizando el conjunto de datos clínicos, validamos que las entradas de múltiples vistas eran beneficiosas para la tarea de clasificación de la densidad mamaria con un aumento de al menos un 2,0% en la precisión y las diferentes vistas conducen a diferentes capacidades de clasificación de modelos.

Proceso o validación:

En este estudio se evaluó dos conjuntos de datos. En primer lugar, se realizó un estudio retrospectivo de mamografías de 1985 recopiladas de 500 pacientes. Los exámenes se realizaron entre octubre de 2015 y agosto de 2018. Las pacientes tenían entre 17 y 75 años. Se toman de 2 a 4 mamografías en diferentes vistas (vistas CC y MLO) para cada paciente. Entre las mamografías de 1985, 319 se clasificaron como BI-RADS I (86 pacientes), 423 fueron BI-RADS II (106 pacientes), 541 fueron BI-RADS III (133 pacientes) y 702 fueron BI-RADS IV (175 pacientes). Todos los exámenes y la clasificación de densidad fueron realizados por radiólogos que tienen décadas de experiencia. Se le da a la paciente una categoría general de BI-RADS sin distinguir la mama izquierda o derecha. Las mamografías se recopilaron y guardaron mediante el formato estándar de comunicaciones e imágenes digitales en médico (DICOM). La matriz de la imagen tiene diferentes tamaños de 407×1374 para 2394×3062 .

Resultado:

las redes convolucionales profundas con aprendizaje residual pueden extraer automáticamente características sutiles, de alto orden y alta abstracción de las mamografías que incluso podrían pasar por alto a simple vista, lo que permite una discriminación precisa de los cuatro BI-RADS categorías. Con la red de aprendizaje residual profunda dilatada y guiada por la atención de entrada única propuesta, una precisión general del 88,7% y 70. El rendimiento de las redes neuronales profundas depende significativamente del conjunto de datos estudiado.

2.1.1.9 Artículo 09 (Dontchos et al., 2020):

Título: Validación externa de un profundo Modelo de aprendizaje para predecir Densidad mamaria mamográfica en Práctica clínica de rutina

Aporte:

El modelo de aprendizaje profundo se implementó en una práctica especializada en imágenes mamarias con personal de radiólogos académicos y comunitarios de imágenes mamarias. La evaluación del modelo de aprendizaje profundo de la densidad mamaria mamográfica se presentó al radiólogo durante la práctica clínica de rutina en el momento de la interpretación de la mamografía. Identificamos 2174 mamografías de detección consecutivas después de la implementación del modelo de aprendizaje profundo. La evaluación mamográfica de la densidad mamaria es subjetiva y varía ampliamente entre y dentro de los radiólogos (4-10). En un estudio de 83 radiólogos, hubo una variación considerable en las evaluaciones cualitativas de densidad del Sistema de datos e informes de imágenes de mama (BI-RADS) con un rango de 6% -85% de mamografías evaluadas como heterogénea o extremadamente densa (4). En otro estudio de 34 radiólogos, el acuerdo de densidad intraradiólogo entre las mujeres con dos exámenes osciló entre el 62% y el 87%.

Proceso o validación:

Siguiendo el léxico BI-RADS del Colegio Americano de Radiología de categoría a (casi completamente graso), b (áreas dispersas de densidad fibroglandular), c (heterogéneamente densa) y d (extremadamente densa). Ocho radiólogos académicos (rango de 2 a 24 años de experiencia; media de 8,3 años; mediana de 4 años) y cinco radiólogos comunitarios (rango de 4 a 20 años de experiencia; media de 11,6 años; mediana de 10 años) interpretaron mamografías de detección. durante la fase de implementación clínica. Los ocho radiólogos académicos y cuatro de los cinco radiólogos comunitarios completaron un programa de capacitación especializado en imágenes de mama (el único radiólogo de la comunidad que no completó una beca dedicada a imágenes de mama ha practicado imágenes clínicas de mama durante 17 años). Ninguno de estos 13 radiólogos contribuyó con evaluaciones prospectivas de densidad al entrenamiento, desarrollo o conjuntos de prueba para el algoritmo DL.

Durante la revisión de la mamografía de detección, todos los radiólogos recibieron la evaluación de densidad del modelo DL en el informe de imágenes electrónico proporcionado

por el software de informes. La evaluación final de la densidad de la mamografía quedó a criterio del radiólogo, es decir, estar de acuerdo o en desacuerdo con el modelo DL. Todas las mamografías se analizaron en nuestras estaciones de trabajo de revisión (Hologic, Bedford, MA) siguiendo nuestro flujo de trabajo clínico de rutina para la evaluación y el informe de mamografías.

Resultado:

Después de la implementación clínica del modelo DL, 2174 mamografías de detección fueron interpretadas de forma prospectiva por radiólogos académicos (1079 exámenes) y comunitarios (1095 exámenes). La edad media, la distribución por edades y la raza fueron similares entre los exámenes evaluados por radiólogos académicos y comunitarios. Los radiólogos académicos evaluaron el 39,3% de los exámenes como densos y los radiólogos comunitarios evaluaron el 42,8% de los exámenes como densos ($p = 0,09$). El modelo DL evaluó una proporción similar de exámenes como densos en los exámenes académicos (34,9%) y comunitarios evaluados por radiólogos.

Los radiólogos tuvieron una alta aceptación clínica de la predicción de densidad del modelo DL, con un 94,9% de aceptación para las categorías densas versus no densas y un 92,1% de aceptación.

2.1.1.10 Artículo 10 (Roth et al., 2020):

Título: Aprendizaje federado para la clasificación de la densidad mamaria: una implementación en el mundo real

Aporte: La construcción de modelos robustos basados en el aprendizaje profundo requiere grandes cantidades de diversos datos de entrenamiento. En este estudio, se investigó el uso de aprendizaje federado (FL) para construir modelos de clasificación de imágenes médicas en un entorno colaborativo del mundo real. Siete instituciones clínicas de todo el mundo se unieron a este esfuerzo de FL para formar un modelo de densidad mamaria clasificación basada en el sistema de datos, informes e imágenes de mama (BIRADS). Muestra que a pesar de las diferencias sustanciales entre los conjuntos de datos de todos los sitios (sistema de mamografía, distribución de clases y conjunto de datos tamaño) y sin centralizar los datos, podemos entrenar con éxito modelos de IA en federación. Los resultados muestran que los modelos entrenados con FL funcionan 6.3% en promedio mejor que sus contrapartes capacitadas en un instituto solo datos locales. Además, mostramos una mejora relativa del

45,8% en la generalizabilidad de los modelos cuando se evalúan en los otros participantes datos de prueba de los sitios.

Proceso o validación:

Utilizando la kappa ponderada lineal de Cohen para evaluar el rendimiento de los mejores modelos locales antes y después del aprendizaje federado en comparación

con las evaluaciones de la verdad sobre el terreno de los radiólogos. La puntuación kappa es un número entre -1 y 1. Las puntuaciones superiores a 0,8 generalmente se consideran muy buenas acuerdo, mientras que cero o menos significaría que no hay acuerdo (prácticamente al azar asignación de etiquetas). Una kappa de 0,21 a 0,40, 0,41 a 0,60 y 0,61 a 0,80 representa acuerdo justo, moderado y sustancial, respectivamente [11]. La kappa La medida ha sido elegida para ser directamente comparable con la literatura previa sobre clasificación de la densidad mamaria en mamografía.

Resultado:

Se muestra el desempeño de los mejores modelos localmente (seleccionados por mejor puntuación de validación en datos locales) utilizando solo datos de entrenamiento local, así como después aprendizaje federado. En promedio, se puede observar una mejora relativa del 6.3% cuando el modelo se aplica a los datos de prueba de un cliente (diag. media). También observamos una mejora general de estos mejores modelos locales aplicados a los diferentes clientes datos de prueba. Aquí, la generalización (media fuera del diagnóstico) de los modelos mejoró por 45,8% de media.

2.1.1.11 Artículo 11 (Maghsoudi et al., 2021):

Título: Deep-LIBRA: un método de inteligencia artificial para la cuantificación sólida de la densidad mamaria con validación independiente en la evaluación del riesgo de cáncer de mama

Aporte: Los autores presentan un método de inteligencia artificial (IA) para estimar la densidad mamaria a partir de mamografías digitales.

Proceso o validación:

El método aprovecha el aprendizaje profundo mediante dos arquitecturas de redes neuronales convolucionales para segmentar con precisión el área de la mama. Luego, se

aplica un algoritmo de inteligencia artificial que combina la generación de superpíxeles y el aprendizaje automático radiómico para diferenciar las regiones de tejido densas de las no densas dentro de la mama, a partir de las cuales se estima la densidad mamaria.

El método fue entrenado y validado en un conjunto de datos multirraciales e institucionales de 15,661 imágenes (4,437 mujeres), y luego probado en un conjunto de datos de casos y controles emparejados de 6368 mamografías digitales (414 casos; 1178 controles) para ambas densidades mamarias se realizó la estimación y discriminación de casos y controles. En el conjunto de datos independientes, las estimaciones de densidad porcentual de mama (PD) de Deep-LIBRA y un lector experto estaban fuertemente correlacionadas (coeficiente de correlación de Spearman = 0,90).

Resultado:

El modelo ajustado por edad e IMC, Deep-LIBRA arrojó un rendimiento de discriminación de casos y controles más alto (área bajo la curva ROC, AUC = 0,612 [intervalo de confianza (IC) del 95%: 0,584, 0,640]) en comparación con otros cuatro Métodos de evaluación de la densidad mamaria comercial y de investigación ampliamente utilizados (AUC = 0,528 a 0,599).

2.1.1.12 Artículo 12 (Chang et al., 2020):

Título: Evaluación multiinstitucional y evaluación de crowdsourcing del aprendizaje profundo para la clasificación automatizada de la densidad mamaria

Aporte:

Los autores proponen el desarrollo de un algoritmo de aprendizaje profundo para evaluar automáticamente la densidad mamaria BI-RADS.

Proceso o validación:

Las mamografías de detección digital de 33 sitios clínicos se obtuvieron retrospectivamente a través del Ensayo de detección de imágenes mamográficas digitales, cuyos detalles se publicaron anteriormente. Cada examen fue interpretado por un solo radiólogo de una cohorte de radiólogos que utiliza el léxico de densidad mamaria ACR BI-RADS (categoría a: grasa, categoría b: dispersa, categoría c: heterogéneamente densa, categoría d: extremadamente densa). Un total de 92 radiólogos leyeron los exámenes. Los lectores en los Estados Unidos eran todos intérpretes calificados de mamografías según la ley federal. Los lectores

canadienses cumplieron con estándares equivalentes. El radiólogo principal de cada sitio recibió capacitación para leer para la Prueba de detección de imágenes mamográficas digitales y, a su vez, capacitó a los demás lectores del sitio. Todas las imágenes fueron desidentificadas previamente antes de este estudio. Las mamografías se guardaron en formato DICOM con cuatro formatos de datos de imagen diferentes, correspondientes a diferentes sistemas de mamografía digital o diferentes versiones del mismo sistema.

Resultado:

Nuestro algoritmo de mejor rendimiento logró un buen acuerdo con los radiólogos que eran intérpretes calificados de mamografías, con un k de 0,667. Cuando el entrenamiento se realizó con imágenes muestreadas al azar del conjunto de datos en lugar de muestrear el mismo número de imágenes de cada categoría de densidad, las predicciones del modelo se desviaron de las categorías de baja prevalencia, como los senos extremadamente densos. El resultado neto fue un aumento en la sensibilidad y una disminución en la especificidad para predecir mamas densas para la misma clase en comparación con el muestreo aleatorio. También encontramos que el desempeño del modelo se degrada cuando evaluamos en formatos de datos de mamografía digital que difieren de aquel en el que nos capacitamos, enfatizando la importancia de los conjuntos de capacitación multiinstitucionales. Por último, mostramos que las anotaciones colaborativas, incluidas las de los asistentes que leen mamografías de forma rutinaria, tenían un mayor acuerdo con nuestro algoritmo que con los radiólogos intérpretes originales.

2.1.1.13 Artículo 13 (Al-Masni et al., 2018):

Título: Detección y clasificación simultánea de masas mamarias en mamografías digitales a través de un sistema CAD de aprendizaje profundo basado en YOLO

Aporte:

En este artículo, los autores proponen un novedoso sistema de diagnóstico asistido por computadora (CAD) basado en una de las técnicas regionales de aprendizaje profundo, una red neuronal convolucional (CNN) basada en ROI que se llama You Only Look Once (YOLO). Aunque la mayoría de los estudios anteriores solo se ocupan de la clasificación de masas, nuestro sistema CAD propuesto basado en YOLO puede manejar la detección y clasificación simultáneamente en un marco. El sistema CAD propuesto contiene cuatro etapas principales: preprocesamiento de mamografías, extracción de características

utilizando redes convolucionales profundas, detección masiva con confianza y, finalmente, clasificación masiva utilizando redes neuronales totalmente conectadas (FC-NN).

Proceso o validación:

El sistema CAD propuesto contiene cuatro etapas principales: preprocesamiento de mamografías, extracción de características utilizando redes convolucionales profundas, detección masiva con confianza y, finalmente, clasificación masiva utilizando redes neuronales totalmente conectadas (FC-NN). En este estudio, utilizamos 600 mamografías originales de la base de datos digital para mamografías de detección (DDSM) y sus mamografías aumentadas de 2400 con la información de las masas y sus tipos en el entrenamiento y prueba de nuestro CAD. El sistema CAD capacitado basado en YOLO detecta las masas y luego clasifica sus tipos en benignos o malignos.

You Only Look Once (YOLO) es una de las técnicas de aprendizaje profundo más modernas [15]. Es capaz de detectar y clasificar objetos en todas las imágenes al mismo tiempo. A diferencia de las técnicas de detección anteriores que aplicaban el clasificador a múltiples regiones de la imagen [8], YOLO utiliza una única red neuronal convolucional para toda la imagen. Este enfoque divide la imagen de entrada en subregiones y predice múltiples cuadros delimitadores con sus probabilidades de clase para cada región. A diferencia de la R-CNN tradicional que requiere muchas redes para todas las regiones extraídas, YOLO utiliza las mamografías completas para que la información contextual de los predictores y su aspecto estén completamente codificados con una sola red tanto en el tiempo de entrenamiento como de prueba [15]. YOLO tiene varias ventajas sobre otros sistemas de detección.

Resultado:

Nuestros resultados con pruebas de validación cruzada de cinco veces muestran que el sistema CAD propuesto detecta la ubicación de la masa con una precisión general del 99,7%. El sistema también distingue entre lesiones benignas y malignas con una precisión global del 97%.

2.1.1.14 Artículo 14 (Tajerian et al., 2021):

Título: Artemisia: validación de un modelo de aprendizaje profundo para la categorización automática de la densidad mamaria

Aporte:

En este artículo presentamos una CNN que había sido capacitada en base a la evaluación cualitativa de la densidad mamaria, cuyo desempeño se comparó con seis expertos en imágenes mamarias. El objetivo de este estudio es validar un algoritmo de aprendizaje profundo, un desarrollo interno, para la clasificación de la densidad mamaria según los patrones de densidad mamaria de ACR. Presumimos que nuestro modelo de aprendizaje profundo alcanzaría un rendimiento de nivel profesional.

Proceso o validación:

Para cada una de las etapas, evaluamos las coincidencias de Artemisia, los observadores, una aplicación de software comercial y el informe mayoritario, mediante el cálculo de coeficientes Kappa ponderados linealmente (k). A su vez, reportamos la variabilidad intraobservador para cada uno de los participantes, comparando la concordancia entre ambas etapas. Para el cálculo utilizamos el método descrito por Cohen y Fleiss (26,27) y tomamos como referencia la subdivisión del coeficiente k de Landis y Koch (0: “pobre”; de 0 a 0,2: “leve”; de 0,21 a 0,4: “regular”; de 0,41 a 0,6: “moderado”; de 0,61 a 0,8: “sustancial”; de 0,81 a 1: “casi perfecto”) (28). Además, según la relevancia clínica, los resultados se dicotomizaron en patrón no denso (A o B) y patrón denso (C o D). El rendimiento diagnóstico de Artemisia también se evaluó en comparación con el informe de la mayoría a través de la sensibilidad, especificidad, valor predictivo positivo y negativo con sus respectivos intervalos de confianza. Elegimos informar estas últimas métricas solo desde la primera etapa. Por un lado, para evitar la redundancia y por otro lado porque reflejan el primer encuentro de los profesionales con las imágenes

Resultado:

Tomamos las categorías asignadas por los seis profesionales, la aplicación de software comercial, la CNN y el informe mayoritario para las 451 imágenes mamográficas. La prevalencia de alta densidad mamaria en la muestra fue del 41% de coeficientes liner kappa presenta las concordancias entre todos los participantes en cada etapa. Los niveles de acuerdo entre la CNN y el informe de la mayoría fueron $k = 0,64$ (IC del 95%: 0,58-0,69) en la primera etapa y $k = 0,57$ (IC del 95%: 0,52-0,63) en la segunda etapa. A su vez, los valores de los profesionales con respecto a la mayoría reportada estuvieron en los rangos $k = (0.64; 0.84)$ en la primera etapa y $k = (0.54; 0.83)$ en la segunda etapa. En cuanto a la aplicación de

software comercial con respecto al reporte mayoritario, la kappa para las dos etapas fueron 0,46 y 0,44, respectivamente. Considerando la categoría dicotomizada denso y no denso, la concordancia entre la CNN y el informe de la mayoría fue $k = 0,71$ (0,64–0,78) y $k = 0,70$ (IC del 95%: 0,63–0,76).

2.1.1.15 Artículo 15 (Salama, 2021):

Título: Aprendizaje profundo en segmentación y clasificación de imágenes de mamografía: enfoque de CNN automatizado

Aporte:

En este trabajo, se propone un nuevo marco para la segmentación y clasificación de imágenes de cáncer de mama.

Proceso o validación:

Se aplican diferentes modelos, incluidos InceptionV3, DenseNet121, ResNet50, VGG16 y Mobile-NetV2, para clasificar la Sociedad de análisis de imágenes mamográficas (MIAS), la Base de datos digital para mamografías de detección (DDSM) y el Subconjunto de imágenes mamarias curadas de DDSM (CBIS-DDSM) en benigno y maligno. Además, el modelo U-Net modificado entrenado se utiliza para segmentar el área del seno a partir de las imágenes de la mamografía. Este método ayudará como asistente de un radiólogo en la detección temprana y mejorará la eficiencia de nuestro sistema. La visión Cranio Caudal (CC) y la Visión Medio lateral Oblicua (MLO) se utilizan ampliamente para la identificación y el diagnóstico del cáncer de mama. La precisión del diagnóstico de cáncer de mama mejorará a medida que aumente el número de visualizaciones. Nuestro marco de trabajo propuesto se basa en la vista MLO y la vista CC para mejorar el rendimiento del sistema. Además, la falta de datos etiquetados es un gran desafío. El aprendizaje por transferencia y el aumento de datos se aplican para superar este problema. Tres conjuntos de datos mamográficos; MIAS, DDSM y CBIS-DDSM, se utilizan en nuestra evaluación. En este artículo se presentan las redes neuronales totalmente convolucionales (CNN) de extremo a extremo. La técnica propuesta de aplicar el aumento de datos con el modelo U-Net modificado e InceptionV3 logra el mejor resultado, específicamente con el conjunto de datos DDSM. Esto logra un 98,87% de precisión, un 98,88% de área bajo la curva (AUC), un 98,98% de sensibilidad, un 98,79% de precisión, un 97,99% de puntuación F1 y un tiempo de cálculo de 1,2134 s en conjuntos de datos DDSM. la falta de datos etiquetados es un gran

desafío. El aprendizaje por transferencia y el aumento de datos se aplican para superar este problema. Tres conjuntos de datos mamográficos; MIAS, DDSM y CBIS-DDSM, se utilizan en nuestra evaluación.

Resultado:

En este artículo se presentan las redes neuronales totalmente convolucionales (CNN) de extremo a extremo. La técnica propuesta de aplicar el aumento de datos con el modelo U-Net modificado e InceptionV3 logra el mejor resultado, específicamente con el conjunto de datos DDSM. Esto logra un 98,87% de precisión, un 98,88% de área bajo la curva (AUC), un 98,98% de sensibilidad, un 98,79% de precisión, un 97,99% de puntuación F1 y un tiempo de cálculo de 1,2134 s en conjuntos de datos DDSM. la falta de datos etiquetados es un gran desafío. El aprendizaje por transferencia y el aumento de datos se aplican para superar este problema. Tres conjuntos de datos mamográficos; MIAS, DDSM y CBISDDSM, se utilizan en nuestra evaluación. En este artículo se presentan las redes neuronales totalmente convolucionales (CNN) de extremo a extremo. La técnica propuesta de aplicar el aumento de datos con el modelo U-Net modificado e InceptionV3 logra el mejor resultado, específicamente con el conjunto de datos DDSM. Esto logra un 98,87% de precisión, un 98,88% de área bajo la curva (AUC), un 98,98% de sensibilidad, un 98,79% de precisión, un 97,99% de puntuación F1 y un tiempo de cálculo de 1,2134 s en conjuntos de datos DDSM. En este artículo se presentan las redes neuronales totalmente convolucionales (CNN) de extremo a extremo. La técnica propuesta de aplicar el aumento de datos con el modelo U-Net modificado e InceptionV3 logra el mejor resultado, específicamente con el conjunto de datos DDSM. Esto logra un 98,87% de precisión de área b.

2.1.1.16 Artículo 16 (Rampun et al., 2020):

Título: Clasificación de densidad mamaria en mamografías: una investigación de técnicas de codificación en patrones locales basados en binarios

Aporte:

Los investigadores proponen un enfoque de codificación local binaria, ternaria y quinario para aplicarlas a la clasificación de densidad mamaria en mamografías. Posteriormente, proponemos un nuevo enfoque de codificación basado en una técnica de siete codificaciones,

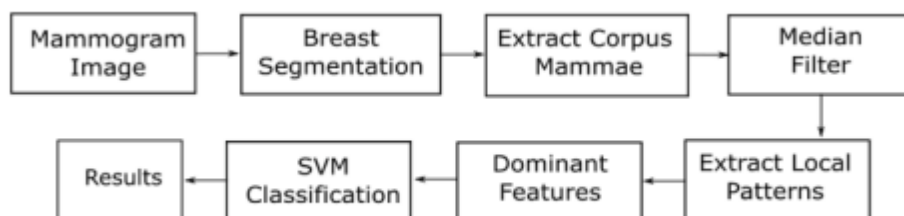
lo que produce un nuevo operador de patrón local llamado operador de patrón septenario local

Proceso o validación:

Esta sección explica los detalles técnicos de nuestro estudio, cubriendo los pasos involucrados en el preprocesamiento, extracción de características y las fases de clasificación, la cual muestra una descripción general del flujo de trabajo en nuestros experimentos.

Figura 4

Flujo de Trabajo



Nota. La imagen muestra el flujo de trabajo involucrado en el experimento al 21 de enero del 2023. Adaptado de “Clasificación de densidad mamaria en mamografías: una investigación de técnicas de codificación en patrones locales basados en binarios”, por Rampun A., 2020

Resultado:

Los resultados experimentales sugieren que el operador de patrón local propuesto es robusto y supera a las otras técnicas de codificación cuando se evalúa en los conjuntos de datos de Mammographic Image Analysis Society (MIAS) e InBreast. El operador de patrón septenario local logró una precisión de clasificación máxima de 83,3% y 80,5% en los conjuntos de datos MIAS e InBreast, respectivamente. La comparación más cercana lograda por los otros operadores de patrones locales es el operador quinario local, con precisiones máximas de 82.

2.1.1.17 Artículo 17 (Saffari et al., 2020):

Título: Segmentación y clasificación de densidad mamaria totalmente automatizada mediante aprendizaje profundo

Aporte:

Este estudio se pretende desarrollar una segmentación y clasificación del tejido mamario totalmente automatizada y digitalizada utilizando técnicas avanzadas de aprendizaje profundo. La red condicional Generative Adversarial Networks (cGAN) se aplica para segmentar los tejidos densos en mamografías. Tener un sistema completo para la clasificación de la densidad mamaria, proponemos una red neuronal convolucional (CNN) para clasificar las mamografías en función de la estandarización del sistema de datos e informes de imágenes de mama (BI-RADS).

Las características novedosas y ventajosas del método propuesto en este estudio son las siguientes:

- Se desarrolla la primera adaptación de cGAN en el área de segmentación de densidad mamaria totalmente automatizada en mamografías,
- la clasificación porcentual de densidad mamaria por la arquitectura CNN de clases múltiples desarrollada se correlacionó bien con las calificaciones de densidad BI-RADS (BI-RADS I, BI-RADS II, BI-RADS III y BI-RADS IV) utilizando la máscara binaria segmentada en la etapa anterior (salida cGAN) por radiólogos,
- Se puede obtener una fuerte correspondencia entre la salida de nuestro algoritmo automatizado y las medidas de densidad mamaria presentadas por el radiólogo, y
- el enfoque propuesto da como resultado un cálculo notablemente más rápido al tiempo que mejora la eficiencia de clasificación en comparación con otros métodos en la literatura.

Proceso o validación:

La red de clasificación se alimenta de las máscaras segmentadas de tejidos densos generadas por la red cGAN. Para la mamografía de detección, se utilizaron 410 imágenes de 115 pacientes del conjunto de datos de INbreast. El marco propuesto puede segmentar las regiones densas con una precisión, coeficiente de datos, índice de Jaccard del 98%, 88% y 78%, respectivamente.

En la validación cruzada reservada, (33%) de los datos pertenecen a BI-RADS I con 108 imágenes de 27 pacientes, (35%) de los datos pertenecen a BI-RADS V con 116 imágenes de 29 pacientes, (25%) de los datos pertenecen a BI-RADS III con 80 imágenes de 20 pacientes, y solo (7%) de los datos pertenecen a BI-RADS IV con 22 imágenes de 6 pacientes. Se muestra la distribución del conjunto de datos de INbreast para los conjuntos de

entrenamiento y prueba. Uno de los principales objetivos de este estudio es crear un modelo robusto que generalice bien a los nuevos datos y utilice imágenes de pacientes no incluidos inicialmente en la etapa de entrenamiento. Para las pruebas, se han utilizado 33 pacientes como un conjunto de control / prueba para evaluar el rendimiento de los modelos profundos entrenados para la segmentación de tejido denso y la clasificación de la densidad mamaria. Por lo tanto, nuestro conjunto de control / prueba puede servir como proxy para nuevos datos.

Resultado:

Entre los resultados se obtuvo precisión, sensibilidad y especificidad de 97,85%, 97,85% y 99,28%, respectivamente, para la clasificación de densidad mamaria. Los hallazgos de este estudio son prometedores y muestran que las técnicas propuestas basadas en el aprendizaje profundo pueden producir una herramienta asistida por computadora clínicamente útil para el análisis de densidad mamaria mediante mamografía digital.

2.1.1.18 Artículo 18 (Lizzi et al., 2019):

Título: Redes neuronales convolucionales residuales para la clasificación de densidad mamaria

Aporte:

Los autores proponen una red neuronal convolucional residual para realizar la clasificación BI-RADS.

Proceso o validación:

El conjunto de datos se ha dividido aleatoriamente en conjunto de formación (1170 exámenes), conjunto de validación (150 exámenes) y conjunto de prueba (150 exámenes). Dado que la densidad mamaria es una evaluación general de las proyecciones, si se produce una asimetría de densidad entre la mama izquierda y la derecha, el radiólogo asigna la clase superior de ese sujeto. Para reproducir tal comportamiento, los puntajes de clasificación se promediaron por separado para el seno derecho e izquierdo y, si ocurría asimetría, se asignó la clase más alta a la mujer.

Resultado:

En ese caso obtuvieron una precisión de prueba igual al 84,9% que es inferior a la alcanzada en este trabajo. En cuanto a la clasificación BIRADS, obtuvimos una precisión de prueba en

imágenes de 650x650 píxeles igual al 78.0%. Este resultado es comparable con el logrado por trabajos anteriores.

2.1.1.19 Artículo 19 (Gorgi et al., 2021):

Título: Validación de un nuevo software totalmente automatizado para la evaluación de la densidad mamaria mamográfica digital 2D para predecir el riesgo de cáncer de mama

Aporte:

Comparamos la precisión para la estratificación del riesgo de cáncer de mama (BC) de un nuevo sistema completamente automatizado (DenSeeMammo - DSM) para la evaluación de la densidad mamaria (BD) con un umbral de no inferioridad basado en la evaluación visual de los radiólogos.

Proceso o validación:

Se realizó un análisis agrupado de 14 267 mamografías 2D recopiladas de mujeres de 48 a 55 años que se sometieron a exámenes de detección de cáncer de mama en tres estudios: RETomo, estudio de Florence y PROCAS. BD se expresó a través de la clasificación de densidad del Sistema de datos e informes de imágenes de mama clínica (BI-RADS). Las mujeres en la categoría BI-RADS D tenían un riesgo 2,6 (IC del 95%: 1,5 a 4,4) y 3,6 (IC del 95%: 1,4 a 9,3) veces mayor de cáncer incidente y de intervalo, respectivamente, que las mujeres en las dos categorías más bajas de TB. La capacidad del DSM para predecir el riesgo de incidencia de cáncer no fue inferior a la evaluación visual de los radiólogos, tanto en el punto estimado como en el límite inferior del IC del 95% (AUC 0,589; IC del 95%: 0,580-0,597) estaban por encima del umbral de evaluación visual predefinido (AUC 0,571). El AUC para los cánceres de intervalo (AUC 0,631; IC del 95%: 0,623–0,639) fue incluso mayor. La BD evaluado con un nuevo método totalmente automatizado se asocia positivamente con el riesgo de BC y no es inferior a la evaluación visual de los radiólogos. Es un marcador aún más fuerte de cáncer de intervalo, lo que confirma un efecto de enmascaramiento apreciable de BD que reduce la sensibilidad de la mamografía.

Resultado:

En total, 14 267 mamografías del mismo número de mujeres, de 48 a 55 años de edad, estaban disponibles para la evaluación de la densidad, de las cuales 322 tenían cáncer (115

en RETomo, 63 en el estudio de Florence y 144 en PROCAS) (Tabla 1). La edad media (DE) de las mujeres en el análisis agrupado fue de $51,0 \pm 1,9$ años, mientras que en RETomo fue de $51,2 \pm 2,1$, en el estudio de Florence $50,8 \pm 1,4$ y en PROCAS $51,3 \pm 2,0$. De 322 cánceres diagnosticados, 98 fueron casos incidentes redondos (74 en el estudio RETomo y 24 en el estudio de Florence), de los cuales 35 fueron cánceres de intervalo (26 en el estudio RETomo y 9 en el estudio de Florence) (Tabla 1). La distribución de BD fue similar en todos los estudios, con un 7,4% en BI-RADS A y un 11,6% en BI-RADS D (tabla1). En RETomo, las mujeres menores de 50 años fueron referidas a un intervalo de 1 año, mientras que las mayores de 50 años fueron referidas a 2 años; BIRADS D fue más frecuente en mujeres menores de 50 años 15,9% vs. 11,1% respecto a los mayores de 50 años (Cuadro Complementario S2).

2.1.1.20 Artículo 20 (Chan, 2019):

Título: Aprendizaje profundo para la evaluación de la densidad mamaria mamografía y más allá

Aporte:

Los autores describen el proceso de entrenamiento y validación de un modelo DL y demuestran el potencial del modelo de aprendizaje automático en una aplicación del mundo real. Lehman et al., (2018) entrenaron el modelo DL con miles de mamografías leídas por 12 radiólogos en su práctica clínica, con la densidad BI-RADS evaluada subjetivamente solo por el radiólogo interpretador como la “verdad” de referencia de cada mamografía.

Proceso o validación:

El modelo DL en el estudio de Lehman et al., (2018) se entrenó con más de 41000 mamografías, solo usó imágenes de "presentación" de un proveedor de un centro académico. Se sabe que la apariencia de una mamografía depende de las técnicas de procesamiento de imágenes, que también pueden afectar la información de la imagen que perciben los radiólogos o la visión por computadora. Para entrenar un algoritmo de aprendizaje automático, generalmente se necesitan grandes muestras representativas de la población de pacientes, incluidas las propiedades de los sistemas de adquisición y procesamiento de imágenes, a los que se pretende aplicar

Resultado:

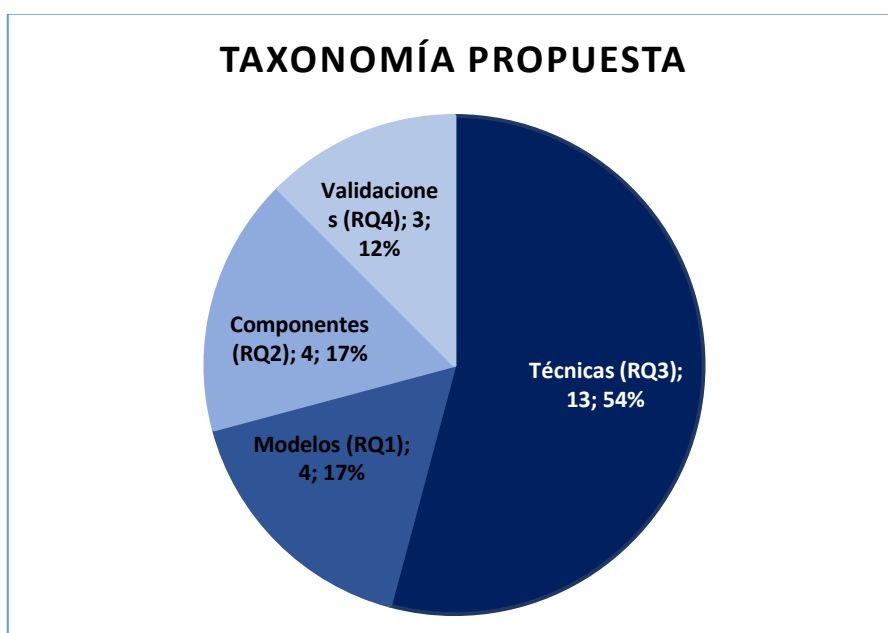
A pesar de estos problemas, que justifican estudios de seguimiento, el estudio de Lehman et al (4), así como muchos otros estudios en la literatura, han demostrado el potencial de la DL para acercar las herramientas de apoyo a la toma de decisiones asistidas por computadora un paso más hacia la práctica clínica. Si la investigación y el desarrollo continúan, podemos esperar que el aprendizaje automático y la inteligencia artificial se conviertan en una parte integral en muchas áreas del proceso de atención médica

2.2 Análisis

Para el análisis de los artículos seleccionados se aplicará una taxonomía desarrollada en función a las preguntas de investigación propuesta en el punto 1.1.

Figura 5

Gráfico de la taxonomía propuesta



La siguiente tabla muestra los artículos seleccionados clasificados según la taxonomía propuesta.

Tabla 4

Clasificación de los artículos según la taxonomía propuesta.

Clasificación	ScienceDirect	Scopus	IEEE	Total
Modelos (RQ1)	[P11, P13]	-	[P02, P10]	4

Componentes (RQ2)	[P11, P13]	-	[P02, P10]	4
Técnicas (RQ3)	[P01, P03, P07, P12, P15, P16, P18]	[P05, P06, P17, P20]	[P04, P08]	13
Validación (RQ4)	[P09]	[P14, P19]	-	3

2.2.1 Modelos (RQ1)

Q1: ¿Qué modelos de clasificación de densidad mamaria existen?

Esta clasificación nos dirá qué temas han cubierto las diversas propuestas en la elección de los modelos de clasificación de densidad mamaria existentes (Q1). Esto requiere una taxonomía basada en los algoritmos de Inteligencia Artificial (IA).

Tabla 5

Modelos para la realización de la clasificación de densidad

Id	Modelos	Fuente
M01	SE-Attention	Deng et al., (2019)
M02	Aprendizaje federado	Roth et al., (2020)
M03	Deep-LIBRA	Maghsoudi et al., (2021)
M04	Yolo	Al-masni et al., (2018)

En cuanto a los modelos que tienen los sistemas de clasificación de densidad mamaria se obtuvo que:

Deng et al., (2019), nos dice que los resultados experimentales demuestran que el marco propuesto obtiene una mayor precisión de clasificación que la red original, como Inception-V4, ResNeXt, DenseNet, aumentando el rendimiento del 89,97% al 92,17%, del 89,64% al 91,57%, del 89,20% al 91,79% respectivamente. Entre ellos, Inception-V4 mejorado posee la mayor precisión, mientras que DenseNet mejora en la mayor medida, tanto el método original como el mejorado.

Según Roth et al., (2020), Se muestra el desempeño de los mejores modelos localmente (seleccionados por mejor puntuación de validación en datos locales) utilizando solo datos de entrenamiento local, así como después aprendizaje federado. En promedio, se puede observar una mejora relativa del 6.3% cuando el modelo se aplica a los datos de prueba de un cliente (diag. media). También observamos una mejora general de estos mejores modelos locales aplicados a los diferentes clientes datos de prueba. Aquí, la generalización (media fuera del diagnóstico) de los modelos mejoró por 45,8% de media. Se resume los puntajes kappa para el entrenamiento local, después de FL, incluyendo después del ajuste fino local, que mejora el rendimiento de un modelo dado en el los datos de prueba locales del cliente en todos menos uno.

Así mismo, Al-masni et al., (2018), menciona que los resultados con pruebas de validación cruzada de cinco veces muestran que el sistema CAD propuesto detecta la ubicación de la masa con una precisión general del 99,7%. El sistema también distingue entre lesiones benignas y malignas con una precisión global del 97%.

Por lo tanto, luego de analizar los papers se evidencia que el uso de los modelos permitirá seguir un patrón para clasificar las mamografías densas con precisión, asimismo se espera reducir costos y eliminar la carga laboral de los radiólogos.

2.2.2 Componentes (RQ2)

Q2: ¿Qué componentes tienen los modelos de clasificación de densidad mamaria?

Esta clasificación nos dirá qué temas han cubierto las diversas propuestas en la elección de los componentes usados para la clasificación de densidad mamaria existentes (Q2). Esto requiere una taxonomía basada en la segmentación y clasificación.

Tabla 6

Componentes para la clasificación de densidad mamaria

Id	Componentes	Fuente
C01	Normalización de imágenes, Entrenamiento, recalibración	Deng et al., (2019)
C02	Sistemas de datos, informes e imágenes mamográficas	Roth et al., (2020)
C03	Arquitectura	Maghsoudi et al., (2021)
C04	Imágenes mamográficas	Al-masni et al., (2018)

En cuanto a los componentes que tienen los sistemas de clasificación de densidad mamaria se obtuvo que:

Deng et al., (2019), Los resultados experimentales demuestran que el marco propuesto obtiene una mayor precisión de clasificación que la red original, como Inception-V4, ResNeXt, DenseNet, aumentando el rendimiento del 89,97% al 92,17%, del 89,64% al 91,57%, del 89,20% al 91,79% respectivamente. Entre ellos, Inception-V4 mejorado posee la mayor precisión, mientras que DenseNet mejora en la mayor medida, tanto el método original como el mejorado.

Según Roth et al., (2020) se muestra el desempeño de los mejores modelos localmente (seleccionados por mejor puntuación de validación en datos locales) utilizando solo datos de entrenamiento local, así como después aprendizaje federado. En promedio, se puede observar una mejora relativa del 6.3% cuando el modelo se aplica a los datos de prueba de un cliente (diag. media). También observamos una mejora general de estos mejores modelos locales aplicados a los diferentes clientes datos de prueba. Aquí, la generalización (media fuera del diagnóstico) de los modelos mejoró por 45,8% de media.

Según Maghsoudi et al., (2021), nos dice que el modelo ajustado por edad e IMC, Deep-LIBRA arrojó un rendimiento de discriminación de casos y controles más alto (área bajo la curva ROC, AUC = 0,612 [intervalo de confianza (IC) del 95%: 0,584, 0,640]) en comparación con otros cuatro Métodos de evaluación de la densidad mamaria comercial y de investigación ampliamente utilizados (AUC = 0,528 a 0,599).

Por lo tanto, luego de analizar los papers se evidencia que los actuales componentes utilizados para la clasificación de mamografías permiten diseñar un sistema de clasificación que pueda reducir la incidencia del cáncer de mama en una etapa temprana.

2.2.3 Técnicas (RQ3)

Q3: ¿Cuáles son las técnicas que se pueden usar para construir estos componentes de clasificación?

Esta clasificación nos dirá qué temas han cubierto las diversas propuestas en la elección de las técnicas usadas en la clasificación de densidad mamaria (Q3). Esto requiere una taxonomía basada en red neuronal convolucional, residuales, múltiples y aprendizaje profundo.

Tabla 7

Técnicas para la realización de la clasificación de densidad mamaria

Id	Técnicas	Fuente
T01	Red neuronal convolucional	Agarwal et al., (2019)
T02	Red neuronal residual profunda	Lizzi et al., (2019)
T03	Aprendizaje profundo	Sun et al., (2019)
T04	Red neuronal convolucional	Trivizakis et al., (2019)
T05	Aprendizaje por transferencia	Sadad et al., (2021)
T06	Aprendizaje profundo	Pérez-Benito et al., (2020)
T07	Aprendizaje residual dilatado	Li et al., (2020)
T08	Aprendizaje profundo	Chang et al., (2020)
T09	Codificación en patrones binarios	Rampun et al., (2020)
T10	Aprendizaje profundo	Saffari et al., (2020)
T11	Redes neuronales convolucionales residuales	Lizzi et al., (2019)
T12	Aprendizaje profundo	Chan (2019)
T13	Deep Learning	Salama (2021)

En cuanto a la mejor técnica para el sistema de clasificación de densidad mamaria se obtuvo que:

Agarwal et al., (2019), nos dice que los resultados demuestran que la técnica propuesta es muy prometedora para ayudar a los radiólogos y servirles como un segundo ojo para clasificar las categorías de densidad mamaria en la detección del cáncer de mama. Los resultados obtenidos de un modelo CNN de ajuste fino convencional se compararon con los

del método propuesto. Los resultados demuestran que la técnica propuesta es muy prometedora para ayudar a los radiólogos y servirles como un segundo ojo para clasificar las categorías de densidad mamaria en la detección del cáncer de mama. Los resultados obtenidos de un modelo CNN de ajuste fino convencional se compararon con los del método propuesto. Los resultados demuestran que la técnica propuesta es muy prometedora para ayudar a los radiólogos y servirles como un segundo ojo para clasificar las categorías de densidad mamaria en la detección del cáncer de mama. La precisión general alcanzó el 88,3% para el método.

Según Lizzi et al., (2019), respecto al problema denso / no denso, la CNN entrenó en 650×650 píxeles imágenes, predice la etiqueta correcta con una precisión igual al 89,4%, que es la mejor exactitud de la prueba obtenida en esta tarea. En cuanto a la clasificación BI-RADS, se obtuvo una exactitud de la prueba en imágenes de 650×650 píxeles igual al 77,3%.

Según Trivizakis et al., (2019), nos dice que todos los modelos se ajustaron en el mismo conjunto de convergencia estratificada (entrenamiento / validación) y se evaluaron en idénticos probar pliegues de validación cruzada para garantizar una comparación justa y transparente. Esto resultó en un 64,6% de formación, un 15,4% de validación y un 20% de pruebas de imágenes de mamografía del mini-MIAS y un 63,9% de formación, un 16,1% de validación y un 20% de pruebas de la base de datos DDSM respectivamente. Se realizaron diferentes algoritmos y estrategias de anotación en los dos conjuntos de datos estudiados para identificar la representación óptima del espacio de características. Las precisiones en el conjunto de datos mini- MIAS variaron del 50,9% (características seleccionadas de GABOR + LBP) al 74,2% (LBP) para la clasificación de tres clases, mientras que, para la clasificación binaria, las puntuaciones de AUC variaron del 48,7% (características seleccionadas de HOG) a 78,0% (LBP). De manera similar, la metodología descrita anteriormente se aplicó en el conjunto de datos DDSM completo para predecir la puntuación MBD en un esquema de anotación binario y de clases múltiples. La puntuación óptima (ACC 79,3% y AUC 84,2%) para las técnicas basadas en características se observó en el análisis binario (mamografías densas o no densas) con el método SURF.

Así mismo, Sadad et al., (2021), indica que el modelo de aprendizaje profundo es eficiente en la clasificación porque no necesita extracción y selección de características. Simplemente

podemos proporcionar imágenes como entrada al modelo en características artesanales proporcionadas a clasificadores de aprendizaje automático convencionales. El marco propuesto puede categorizar las imágenes de mamografía en clases densas.

Por lo tanto, luego de analizar los papers se evidencia que existen diversas técnicas que contienen componentes para clasificar la densidad mamaria en mamografías, pero de acuerdo a los resultados la técnica de Redes Neuronales Convolucionales cuenta con un resultado de 85,4% en reconocer, de manera asertiva, la densidad mamaria. Con ello se estima prevenir el cáncer de mama.

2.2.4 Validación (RQ4)

Q4: ¿Cómo se valida los modelos de clasificación de densidad mamaria?

Esta clasificación nos dirá qué temas han cubierto las diversas propuestas en la elección de las validaciones de los modelos para la clasificación de densidad mamaria (Q4). Esto requiere una taxonomía basada en validaciones algoritmos de Inteligencia Artificial.

Tabla 8

Validaciones que se realizan en la clasificación de densidad mamaria

Id	Validación	Fuente
V01	Validación de modelo de aprendizaje	Dontchos et al., (2019)
V02	Validación de modelo de aprendizaje profundo	Tajerian et al., (2021)
V03	Validación de la evaluación de la densidad	Giorgi et al., (2021)

En cuanto a cómo se valida el sistema de clasificación de densidad mamaria se obtuvo que:

Dontchos et al., (2019), que después de la implementación clínica del modelo DL, 2174 mamografías de detección fueron interpretadas de forma prospectiva por radiólogos académicos (1079 exámenes) y comunitarios (1095 exámenes). La edad media, la distribución por edades y la raza fueron similares entre los exámenes evaluados por radiólogos académicos y comunitarios. Los radiólogos académicos evaluaron el 39,3% de los exámenes como densos y los radiólogos comunitarios evaluaron el 42,8% de los exámenes como densos ($p = 0,09$). El modelo DL evaluó una proporción similar de exámenes como densos en los exámenes académicos (34,9%) y comunitarios evaluados por radiólogos. Los radiólogos tuvieron una alta aceptación clínica de la predicción de densidad

del modelo DL, con un 94,9% de aceptación para las categorías densas versus no densas y un 92,1% de aceptación.

Según Tajerian et al., (2021), las categorías asignadas por los seis profesionales, la aplicación de software comercial, la CNN y el informe mayoritario para las 451 imágenes mamográficas. La prevalencia de alta densidad mamaria en la muestra fue del 41% de coeficientes liner kappa presenta las concordancias entre todos los participantes en cada etapa. Los niveles de acuerdo entre la CNN y el informe de la mayoría fueron $k = 0,64$ (IC del 95%: 0,58-0,69) en la primera etapa y $k = 0,57$ (IC del 95%: 0,52-0,63) en la segunda etapa. A su vez, los valores de los profesionales con respecto a la mayoría reportada estuvieron en los rangos $k = (0,64; 0,84)$ en la primera etapa y $k = (0,54; 0,83)$ en la segunda etapa. En cuanto a la aplicación de software comercial con respecto al reporte mayoritario, la kappa para las dos etapas fueron 0,46 y 0,44, respectivamente. Considerando la categoría dicotomizada denso y no denso, la concordancia entre la CNN y el informe de la mayoría fue $k = 0,71$ (0,64–0,78) y $k = 0,70$ (IC del 95%: 0,63–0,76).

Así mismo, Giorgi et al., (2021), nos dice que en total 14 267 mamografías del mismo número de mujeres, de 48 a 55 años de edad, estaban disponibles para la evaluación de la densidad, de las cuales 322 tenían cáncer (115 en RETomo, 63 en el estudio de Florence y 144 en PROCAS) (Tabla 1). La edad media (DE) de las mujeres en el análisis agrupado fue de $51,0 \pm 1,9$ años, mientras que en RETomo fue de $51,2 \pm 2,1$, en el estudio de Florence $50,8 \pm 1,4$ y en PROCAS $51,3 \pm 2,0$. De 322 cánceres diagnosticados, 98 fueron casos incidentes redondos (74 en el estudio RETomo y 24 en el estudio de Florence), de los cuales 35 fueron cánceres de intervalo (26 en el estudio RETomo y 9 en el estudio de Florence) (Tabla 1). La distribución de BD fue similar en todos los estudios, con un 7,4% en BI-RADS A y un 11,6% en BI-RADS D (tabla1). En RETomo, las mujeres menores de 50 años fueron referidas a un intervalo de 1 año, mientras que las mayores de 50 años fueron referidas a 2 años; BIRADS D fue más frecuente en mujeres menores de 50 años 15,9% vs. 11,1% respecto a los mayores de 50 años (Cuadro Complementario S2).

Por lo tanto, luego de analizar los papers se evidencia que existen ya proyectos para es predecir si la clasificación de la mamografía es densa o no, aplicando modelos de aprendizaje y con ello tenemos un mejor panorama de lo que queremos desarrollar.

2.2.5 Análisis cruzado

Tabla 9

Análisis cruzado entre modelos y componentes

Modelos Componentes	M01	M02	M03	M04
C01	[P02]			
C02		[P10]		
C03			[P11]	
C04				[P13]

Según Al-Masni et al., (2018), menciona que You Only Look Once (YOLO) es una de las técnicas de aprendizaje profundo más modernas. Es capaz de detectar y clasificar objetos en todas las imágenes al mismo tiempo. A diferencia de las técnicas de detección anteriores que aplicaban el clasificador a múltiples regiones de la imagen [8], YOLO utiliza una única red neuronal convolucional para toda la imagen. Este enfoque divide la imagen de entrada en subregiones y predice múltiples cuadros delimitadores con sus probabilidades de clase para cada región. A diferencia de la R-CNN tradicional que requiere muchas redes para todas las regiones extraídas, YOLO utiliza las mamografías completas para que la información contextual de los predictores y su aspecto estén completamente codificados con una sola red tanto en el tiempo de entrenamiento como de prueba [15]. YOLO tiene varias ventajas sobre otros sistemas de detección.

Por lo que se considera, en ese sentido que el M04 (YOLO), que contiene el C04 (imágenes mamográficas) sería el mejor modelo y componente para tomar como referencia para crear nuestro modelo de clasificación de densidad mamaria, a través de imágenes mamográficas.

Tabla 10

Análisis cruzado entre validaciones y técnicas.

Validaciones Técnica	V01	V02	V03
T01	[P09]		
T02	[P09]		
T03		[P14]	
T04			[P19]

Según Tajerian et al., (2021), tomamos las categorías asignadas por los seis profesionales, la aplicación de software comercial, la CNN y el informe mayoritario para las 451 imágenes mamográficas. La prevalencia de alta densidad mamaria en la muestra fue del 41% de coeficientes liner kappa presenta las concordancias entre todos los participantes en cada etapa. Los niveles de acuerdo entre la CNN y el informe de la mayoría fueron $k = 0,64$ (IC del 95%: 0,58-0,69) en la primera etapa y $k = 0,57$ (IC del 95%: 0,52-0,63) en la segunda etapa. A su vez, los valores de los profesionales con respecto a la mayoría reportada estuvieron en los rangos $k = (0,64; 0,84)$ en la primera etapa y $k = (0,54; 0,83)$ en la segunda etapa. En cuanto a la aplicación de software comercial con respecto al reporte mayoritario, la kappa para las dos etapas fueron 0,46 y 0,44, respectivamente. Considerando la categoría dicotomizada denso y no denso, la concordancia entre la CNN y el informe de la mayoría fue $k = 0,71$ (0,64–0,78) y $k = 0,70$ (IC del 95%: 0,63–0,76).

En tal sentido, luego de obtener resultados referenciales que se realizó gracias a la validación, podemos recalcar que la técnica más precisa o con más alto auc sería las redes neuronales convolucionales, ya que se utilizarán imágenes mamográficas para la clasificación de densidad mamaria.

3 CONCLUSIONES

Según la investigación y los resultados obtenidos, podemos concluir que, gracias a esta realización del estado del arte, se puede tener mayores referencias y puntos de vistas, respecto a cada uno de los investigadores, los cuales nos ayudan a tener ideas más claras e información más objetiva acerca de lo que realmente queremos abarcar y mejorar en la problemática planteada.

Por otro lado, concluimos también que referente a los modelos existentes se ha evidenciado, respecto a la información brindada por los investigadores, que es necesario realizar la mayor cantidad de validaciones cruzadas ya que esto va a ayudar a que la clasificación de la densidad mamaria sea precisa y se pueda alcanzar hasta un 99% de exactitud. En cuanto a los componentes, los autores concluyen, que se deben utilizar para la data técnicas como la normalización en las mamografías para que al momento de entrenar la información no se obtengan resultados incorrectos.

En el caso de las técnicas que nos permiten construir componentes de clasificación, entre las existentes, la gran mayoría de autores coinciden en que la más adecuada para trabajar con mamografías son las redes neuronales convolucionales ya que se puede trabajar desde muy pocas hasta grandes cantidades de imágenes, por ello se concluye que será la utilizada para nuestro proyecto. Para la validación de la técnica, los investigadores concluyen que es adecuado realizarlo con la ayuda de métricas como la exactitud y la precisión las cuales me permiten evidenciar de manera cuantitativa y cualitativa los resultados obtenidos permitiendo compararla con otras.

Finalmente, podemos concluir que de acuerdo a la revisión realizada y al análisis cruzado, la técnica, componente, modelo y validación seleccionada permiten que nuestra propuesta, en la realización de la clasificación de la densidad mamaria, tenga información detallada que luego aportará para que futuras generaciones la usen como guía en la elaboración de sus proyectos de investigación para apoyar a instituciones de salud las cuales cada día se ven obligadas a aplicar herramientas tecnológicas y soluciones para potenciar su productividad.

4 REFERENCIAS

- Al-masni, M., Al-antari, M., Park, J., Gi, G., Kim, T., Rivera, P., Valarezo, E., Choi, M., Han, S., & Kim, T.-S., (2018, abril). *Simultaneous Detection and Classification of Breast Masses in Digital Mammograms via a Deep Learning YOLO-based CAD System*. ScienceDirect. Recuperado el 10 de marzo del 2021 de: <https://www.sciencedirect.com/science/article/abs/pii/S0169260717314980>
- Chan, H. & Helvie, M., (2019, enero). *Deep Learning for Mammographic Breast Density Assessment and Beyond*. Scopus. Recuperado el 11 de marzo del 2021 de : <https://www.scopus.com/record/display.uri?eid=2-s2.0-85058887277&origin=resultslist&sort=plf-f&src=s&st1=Deep+Learning+for+Mammographic+Breast+Density++Assessment+and+Beyond&sid=60859b75e9b3ad93e44a56873c11b4dd&sot=b&sdt=b&sl=83&s=TILE-ABS-KEY%28Deep+Learning+for+Mammographic+Breast+Density++Assessment+and+Beyond%29&relpos=0&citeCnt=4&searchTerm=>
- Chang K., Beers, A., Brink, L., Patel, J., Singh, P., Arun, N., Hoebel, K., Gaw, N., Shah, M., Pisano, E., Tilkin, M., Coombs, L., Dreyer, K., Allen, B., Agarwal, S. & Kalpathy-Cramer, J., (2020, diciembre). *Multi-Institutional Assessment and Crowdsourcing Evaluation of Deep Learning for Automated Classification of Breast Density*. ScienceDirect. Recuperado el 10 de marzo del 2021 de: <https://www.sciencedirect.com/science/article/abs/pii/S1546144020305391>
- Deng, J., Ma, Y., Li, D., Zhao, J., Liu, Y. & Zhang, H., (2019, septiembre). *Classification of breast density categories according to SE-Attention neural networks*. ScienceDirect. Recuperado el 14 de marzo del 2021 de: [https://www.sciencedirect.com/science/article/pii/S0169260719311058#:~:text=A%20new%20benchmarking%20dataset%20was,%2C%20and%20D%20\(dense\)](https://www.sciencedirect.com/science/article/pii/S0169260719311058#:~:text=A%20new%20benchmarking%20dataset%20was,%2C%20and%20D%20(dense))
- Dontchos, B., Meng, A., Barzilay, R., Xiang, J. & Lehman, C., (2020, 20 de febrero). *External Validation of a Deep Learning Model for Predicting Mammographic Breast Density in Routine Clinical Practice*. ScienceDirect. Recuperado el 12 de marzo del 2021 de: [https://www.academicradiology.org/article/S1076-6332\(19\)30626-9/pdf](https://www.academicradiology.org/article/S1076-6332(19)30626-9/pdf)
- Kitchenham, B., & Charters, S., (2007, enero). *Adaptado de Revisión Sistemática de la literatura*. ResearchGate. Recuperado el 21 de enero del 2023 de: https://www.researchgate.net/publication/302924724_Guidelines_for_performing_Systematic_Literature_Reviews_in_Software_Engineering
- Lehman, C., Yala, A., Schuster, T., Dontchos, B., Bahl, M., Swanson, K. & Barzilay, R., (2018, 16 de marzo). *Mammographic Breast Density Assessment Using Deep Learning: Clinical Implementation*. Pubs RSNA. Recuperado el 12 de marzo del 2021 de: <https://pubs.rsna.org/doi/full/10.1148/radiol.2018180694>
- Li, C., Xu, J., Liu, Q., Zhou, Y., Mou, L., Pu, Z., Xia, Y., Zheng, H. & Wang, S., (2020, 03 de febrero). *Multiple View Mammographic Density Classification by Attention-Guided, Dilated Residual Learning*. IEEE. Recuperado el 14 de marzo del 2021 de: <https://ieeexplore.ieee.org/document/8978513>

- Lizzi, F., Laruina, F., Oliva, P., Retico, A., & Fantacci, M., (2019, 23 de agosto). *Residual convolutional neural networks to automatically extract significant breast density characteristics*. ScienceDirect. Recuperado el 13 de marzo del 2021 de: https://link.springer.com/chapter/10.1007/978-3-030-29930-9_3
- Lizzi, F., Atzori, S., Aringhieri, G., Bosco, P., Marini, C., Retico, A., Traino, A., Caramella, D. & Fantacci, E., (2019, 4 de noviembre). *Residual Convolutional Neural Networks for Breast Density Classification*. ScienceDirect. Recuperado el 12 de marzo del 2021 de: <https://pdfs.semanticscholar.org/b870/50db240860be92a2e0b112eaa7a17932f9e7.pdf>
- Matsuyama, E., Takerhara, M. & Tsai, D., (2020). *Use of a wave-based and fitted convolutional neural network for classification of breast density in mammographic images*. ScienceDirect. Recuperado el 13 de marzo del 2021 de: <https://www.scirp.org/journal/paperinformation.aspx?paperid=97911>
- Maghsoudi O., Gastounioti, A., Scott, C., Pantalone, L., Wu, F., Cohen, E., Winham, S., Conant, E., Vachon, C. & Kontos, D., (2021, octubre). *Deep-LIBRA: An artificial-intelligence method for robust quantification of breast density with independent validation in breast cancer risk assessment*. ScienceDirect. Recuperado el 11 de marzo del 2021 de: <https://www.sciencedirect.com/science/article/pii/S1361841521001845>
- Pérez-Benito, F., Signol, F., Perez- Cortes, J., Fuster-Baggetto, A., Pollan, M., Pérez-Gómez, B., Salas-Trejo, D., Casals, M., Martínez, I. & Llobet, R., (2020, octubre). *A deep learning system to obtain optimal parameters for threshold-based breast and dense tissue segmentation*. ScienceDirect. Recuperado el 12 de marzo del 2021 de: <https://www.sciencedirect.com/science/article/pii/S0169260720315017>
- Rampun, A., Morrow, P., Scotney, B. & Wang, H., (2020, Julio). *Breast density classification in mammograms: An investigation of encoding techniques in binary-based local patterns*. ScienceDirect. Recuperado el 13 de marzo del 2021 de: <https://www.sciencedirect.com/science/article/abs/pii/S001048252030202X>
- Giorgi, P., Djuric, O., Hélin, V., Astley, S., Mantellini, P., Nitrosi, A., Harkness, E., Gauthier, E., Puliti, D., Balleyguier, C., Baron, C. & Gilbert, F., (2021, diciembre). *Validation of a new fully automated software for 2D digital mammographic breast density evaluation in predicting breast cancer risk*. Scopus. Recuperado el 14 de marzo del 2021 de: <https://www.scopus.com/record/display.uri?eid=2-s2.0-85116451203&origin=resultslist&sort=plf-f&src=s&st1=Validation+of+a+new+fully++automated+software+for+2D+digital++mammographic+breast+density++evaluation+in+predicting+breast++cancer+risk&sid=a83cf476f0a6ff198b06d8bae705c425&sort=b&sdt=b&sl=150&s=TITLE-ABS-KEY%28Validation+of+a+new+fully++automated+software+for+2D+digital++mammographic+breast+density++evaluation+in+predicting+breast++cancer+risk%29&relpos=0&citeCnt=0&searchTerm=>
- Roth, H., Chang, K., Singh, P., Neumark, N., Li, W., Gupta, V., Gupta, S., Qu, L., Ihsani, A., Bizzo, B., Wen, Y., Buch, V., Shah, M., Kitamura, F., Mendonza, M., Lavor, V., Harouni, A., Compas, C., Tetreault, J., Dogra, P., Cheng, Y., Erdal, S., White, R., Hashemian, B., Schultz, T., Zhang, M., McCarthy, A., Yun, B., Sharaf, E., Hoebel, K.,

- Patel, J., Chen, B., Ko, S., Leibovitz, E., Pisano, E., Coombs, L., Xu, D., Dreyer, K., Dayan, I., Naidu, R., Flores, M., Rubin, D. & Kalpathy-Cramer, J., (2020, 3 de septiembre). *Federated Learning for Breast Density Classification: A Real-World Implementation*. ScienceDirect. Recuperado el 14 de marzo del 2021 de: <https://www.sciencedirect.com/record/display.uri?eid=2-s2.0-85092133556&origin=resultslist&sort=plf-f&src=s&st1=Federated+Learning+for+Breast+Density+Classification%3a+A+Real-World+Implementation&sid=50423a11201fd52c7f390ecae339fb74&sot=b&sdt=b&sl=96&s=TITLE-ABS-KEY%28Federated+Learning+for+Breast+Density+Classification%3a+A+Real-World+Implementation%29&relpos=1&citeCnt=5&searchTerm=>
- Sadad, T., Khan, A., Hussain, A., Tariq, U., Fati, S., Bahaj, S. & Munir, A., (2021, 27 de abril). *Internet de las cosas médicas que integra el aprendizaje profundo con datos aumento para clasificación de densidad de mamografía*. Wiley Analytical Science. Recuperado el 11 de marzo del 2021 de: <https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/epdf/10.1002/jemt.23773>
- Saffari, N., Rashwan, H., Abdel-Nasser, M., Singh, V., Arenas, M., Mangina, E., Herrera, B. & Puig, D., (2020, noviembre). *Fully Automated Breast Density Segmentation and Classification Using Deep Learning*. Scopus. Recuperado el 13 de marzo del 2021 de: <https://www.scopus.com/record/display.uri?eid=2-s2.0-85106299573&origin=resultslist&sort=plf-f&src=s&st1=Fully+Automated+Breast+Density+Segmentation+and+Classification+Using+Deep+Learning&sid=6e1f03a974d954fab78939ce22c9cd7e&sot=b&sdt=b&sl=97&s=TITLE-ABS-KEY%28Fully+Automated+Breast+Density+Segmentation+and+Classification+Using+Deep+Learning%29&relpos=0&citeCnt=4&searchTerm=>
- Salama, W. & Aly, M., (2021, octubre). *Deep Learning in Mammography Image Segmentation and Classification - Automated CNN Approach*. ScienceDirect. Recuperado el 13 de marzo del 2021 de: <https://www.sciencedirect.com/science/article/pii/S1110016821002027>
- Sun, L., Wang, J., Hu, Z., Xu, Y., & Cui, Z., (2019, 03 de septiembre). *Convolutional Neural Networks Multiple Views for Mammographic Image Classification*. IEEE. Recuperado el 13 de marzo del 2021 de: <https://ieeexplore.ieee.org/abstract/document/8822935>
- Tajerian, M., Pesce, K., Frangella, J., Quiroga, E., Boietti, B., Chico, M., Swiecicki, M., Benitez, S., Rabellino, M. & Luna, D., (2021, junio). *Artemisia: validation of a deep learning model for automatic breast density categorization*. Scopus. Recuperado el 12 de marzo del 2021 de: <https://www.scopus.com/record/display.uri?eid=2-s2.0-85118233936&origin=resultslist&sort=plf-f&src=s&st1=Artemisia%3a+validation+of+a+deep+learning+model+for+automatic++breast+density+categorization&sid=672745ecab4492e72abb5a0784adb402&sot=b&sdt=b&sl=106&s=TITLE-ABS-KEY%28Artemisia%3a+validation+of+a+deep+learning+model+for+automatic++breast+density+categorization%29&relpos=0&citeCnt=0&searchTerm=>

- Trivizakis, E., Ioannidis, G., Melissianos, V. & Papadakis, G., (2019, septiembre). *A novel deep learning architecture outperforming 'off-the-shelf' transfer learning and feature-based methods in the automated assessment of mammographic breast density*. ResearchGate. Recuperado el 13 de marzo del 2021 de: https://www.researchgate.net/publication/335782328_A_novel_deep_learning_architecture_outperforming_'off-the-shelf'_transfer_learning_and_feature-based_methods_in_the_automated_assessment_of_mammographic_breast_density
- Wong, L., Mauricio, D. & Rodríguez, G., (2017). *About Software Requirements Elicitation*. SemanticScholar. Recuperado el 21 de enero del 2023 de: <https://www.semanticscholar.org/paper/ABOUT-SOFTWARE-REQUIREMENTS-ELICITATION-Wong-Mauricio/503dad846231da01aecb57ff880402e4d4c2ba38>